الإقتصاد القياسي

الجزء الثاني

تأليف

د. جوجارات

D. Gujaratr



تعريب ومراجعة

أ.م.د. هند عبد الغفار عودة

الاقتصاد القياسي

الاقتصاد القياسي

الجزء الثاني

تأليف

دامودار جيجاراتي

Damodar N. Gujaratic

ترجمة ومراجعة

أ.م.د. هند عبد الغفار عودة

رئيس قسم الإحصاء التطبيقي

كلية التجارة - جامعة حلوان



المملكة العربية السعودية - الرياض - هاتف: 4658523 - 4647531 + (009661)

ص. ب : 10720 - الرمز البريدي : 11443 - فاكس : 4657939 + (009661)

الطبعة الإنجليزية :

BASIC ECONOMETRICS

BY: Damodar N. Gudjratic

ردمك: 6 - 673 - 24 - 9960

المملكة العربية السعودية - الرياض - ص. ب : 10720 - الرمز البريدي : 11443

هاتف : 4647531 / 4658523 فاكس: 4657939 + (009661)

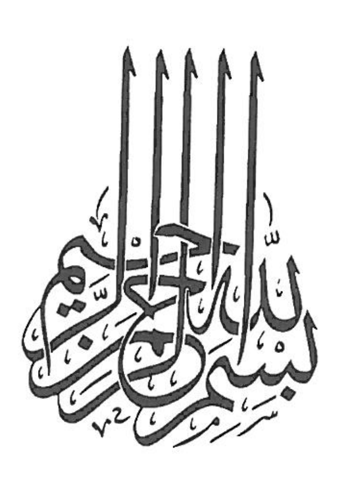
البريد الإلكتروني : Em.il: m.rs@m.rspubl.com

التوزيع داخل جمهورية مصر العربية والسودان وشمال افريقيا :

دار المريخ للنشر بالقاهرة - 4 شارع الفرات - المهندسين - الجيزة - الرمز البريدي: 12411

هاتف : 33376579 / 37609971 فاكس: 37609457 + (00202)

البريد الإلكتروني : Em.il: m.rspub2002@Y.hoo.com


بسم الله الرحمن الرحيم

محتويات الكتاب

(الجزء الأول : من الفصل الأول إلى الفصل الثالث عشر)
(الجزء الثاني : من الفصل الرابع عشر إلى الفصل الثاني والعشرين)

الموضــــوع — الصفحة

الجزء الثالث

موضوعات في الاقتصاد القياسي

الفصل الرابـع عشر

نماذج الانحدار غير الخطية

الفصل الخامس عشر

نماذج الانحدار ذات الاستجابة النوعية

الفصل السادس عشر

نماذج انحدار البيانات

الفصل السابع عشر

نماذج الاقتصاد القياسي الديناميكية

نماذج الانحدار الذاتي . . ونماذج القيم الموزعة متأخراً

الجزء الثالث

نماذج المعادلات الآنية

الفصل الثامن عشر

نماذج المعادلات الآنية

الفصل التاسع عشر

مشكلة التوصيف

الفصل العشرون

طرق المعادلات الآنية

الفصل الحادي والعشرون

السلاسل الزمنية في الاقتصاد القياسي.. بعض المفاهيم الأساسية

الفصل الثاني والعشرون

السلاسل الزمنية في الاقتصاد القياسي: التنبؤ

ملحق A

مراجعة على بعض المفاهيم الإحصائية

ملحق B

مبادئ جبر المصفوفات

ملحق C

طريقة المصفوفات لنماذج الانحدار الخطي

ملحق D

الجزء الثالث

موضوعات في الاقتصاد القياسي

TOPICS IN ECONOMETRICS

في الجزء الأول ، قدمنا نموذج الانحدار الخطي التقليدي مع كل الفروض الخاصة به . وفي الجزء الثاني ، درسنا بالتفصيل العواقب المترتبة على عدم تحقق واحد أو أكثر من هذه الفروض ، وكيفية التعامل مع ذلك . أما في الجزء الثالث ، فإننا سوف ندرس بعض الطرق والأساليب المختارة في الاقتصاد القياسي والمعروفة بتعارضها . وعلى وجه الخصوص سنناقش الموضوعات التالية :

(1) نماذج الانحدار غير الخطية في المعلمات ، (2) نماذج الانحدار ذات الاستجابة النوعية ، (3) نماذج الانحدار ذات البيانات التجميعية و (4) نماذج الاقتصاد القياسي الديناميكية .

وفي الفصل 14 ، سندرس النماذج غير الخطية جوهرياً في المعلمات ، حيث تقدير معالم هذه النماذج لم يعد تحدياً كبيراً مع المتاح الآن من حزم البرامج الإلكترونية ، وعلى الرغم من الصعوبة التي قد يواجهها بعض القراء والمتعلقة بالخلفية الرياضية للموضوع ، فإن الأفكار الأساسية الخاصة بنماذج الانحدار غير الخطي في المعلمات ، سيتم توضيحها بسهولة في هذا الفصل . وباستخدام الأمثلة المناسبة ، سيتم توضيح كيفية تقدير هذه النماذج وتفسيرها .

أما في الفصل 15 ، سندرس نماذج الانحدار التي يكون فيها المتغير التابع متغيرًا نوعيًا بطبيعته . وبذلك يكون هذا الفصل ، مكملاً للفصل 9 ، والذي

تناولنا فيه النماذج التي تكون فيها المتغيرات المفسرة متغيرات نوعية بطبيعتها. النقطة المهمة في هذا الفصل، هي تطوير النماذج التي تكون فيها بيانات المتغير التابع من نوع نعم أو لا.

ونظراً للمشاكل التي تواجه تقدير هذه النماذج بطريقة المربعات الصغرى العادية (*OLS*) فإن هناك بدائل أخرى مقدمة في هذا الفصل، وهما بديلان: نموذج *Logit* ونموذج *probit*، ويتناول هذا الفصل أيضاً أنواعاً عديدة من نماذج الاستجابة النوعية، مثل نموذج *tobit*، ونموذج انحدار بواسون. هذا بالإضافة إلى تناول مختصر لعدد من نماذج الاستجابة النوعية الأخرى، مثل نموذج *Logit* الترتيبي، ونموذج *probit* الترتيبي، ونموذج *Logit* المتعدد الحدود.

في الفصل 16، سندرس نماذج الانحدار ذات البيانات التجميعية. وهذه النماذج يتم فيها دمج مفردات السلاسل الزمنية مع مفردات جداول القطاع العرضي. وعلى الرغم من أن هذا الدمج يزيد من حجم العينة، فإن مشاكل نماذج الانحدار ذات البيانات التجميعية تظل موجودة. وفي هذا الفصل نقدم أساسيات هذه النماذج، ونرشد القارئ إلى مصادر أخرى مناسبة لدراسات أكثر عمقاً.

كذلك سندرس نماذج الانحدار التي توجد فيها قيم للمتغيرات المفسرة في فترة زمنية حالية بالإضافة إلى قيم نفس المتغيرات في فترات زمنية متأخرة أو سابقة. وسندرس أيضاً النماذج التي تستخدم المتغير التابع في فترات زمنية متأخرة، كأحد المتغيرات المفسرة في النموذج. وهذه النماذج يطلق عليها النماذج الموزعة متأخراً ونماذج الانحدار الذاتي بالترتيب.

وعلى الرغم من الأهمية القصوى للجانب التطبيقي لهذه النماذج في الاقتصاد القياسي، فإنها تحتوي على بعض المشاكل الخاصة بالتقدير، حيث إن هذه النماذج تخالف واحداً أو أكثر من فروض نماذج الإنحدار التقليدية. وسندرس هذه المشاكل الخاصة في إطار نموذج *koyck*، ونماذج التوقعات المعدلة (*AE*)، ونماذج التعديلات الجزئية.

ونتناول أيضاً الانتقادات الموجهة لنماذج التوقعات المعدلة والمقدمة من المدافعين عما يقال عنه مدرسة التوقعات الرشيدة (*RE*).

الفصل الرابع عشر

نمـــاذج الانـــحدار غـــير الخطـــية
NONLINEAR REGRESSION MODELS

إن الاهتمام الرئيسي لهذا الكتاب ، ينصب على نماذج الانحدار الخطية ، وهي النماذج التي تكون خطية في المعلمات ، أو التي يمكن تحويلها إلى نماذج خطية في المعلمات ، أو كليهما معاً . في بعض الأحيان ، سواء من الجانب النظري أو التطبيقي نحتاج إلى التعامل مع نماذج غير خطية في المعلمات .[1] في هذا الفصل نركز الاهتمام على مثل هذه النماذج ، ودراسة الصفات الخاصة بها .

1.14 نماذج الانحدار الخطية جوهرياً وغير الخطية جوهرياً :
INTRINSICALLY LINEAR AND INTRINSICALLY NON LINEAR REGRESSION MODELS

عندما بدأنا الحديث عن نماذج الانحدار الخطية في الفصل الثاني ، وضحنا أن همنا الرئيسي في هذا الكتاب ، ينصب بشكل رئيسي على النماذج الخطية في المعلمات ، والتي قد تكون خطية أو غير خطية في المتغيرات . بالعودة إلى جدول (2.3) تستطيع أن ترى أن النماذج الخطية هي إما نماذج خطية في المعلمات والمتغيرات معاً ، أو خطية في المعلمات وغير خطية في المتغيرات . وعلى الجانب الآخر ، إذا كان

---

(1) لاحظنا في الفصل (4) ، أنه عند تحقق فرض التوزيع الطبيعي لمقدار الخطأ الموجود في النموذج ، فإن مقدرات OLS ليست فقط أفضل مقدرات غير متحيزة خطية BLUE ، وإنما هي أيضاً أفضل مقدرات غير متحيزة BOE من بين كل المقدرات سواء خطية أو غير خطية . ولكن إذا أسقطنا فرض التوزيع الطبيعي ، كما في Hackinnon و Davidson فمن الممكن الحصول على مقدرات غير خطية أو متحيزة تكون أفضل من مقدرات المربعات الصغرى ، انظر :

Russell Davidson and James G. MacKinnon, Estimation and Inference in Econometrics, Oxford University Press, New York, 1993, p. 131.

النموذج غير خطي في المعلمات ، فإنه يكون نموذج انحدار غير خطي ، حتى ولو كان خطياً أو غير خطي في المتغيرات .

وهنا لابد أن نأخذ في الاعتبار نقطة مهمة ، فهناك بعض النماذج التي تظهر وكأنها غير خطية في المعلمات ، رغم أنها جوهرياً هي نماذج خطية ، والدليل على ذلك ، أنه باستخدام تحويلات مناسبة ممكن جعلها نماذج انحدار خطية ، ولكن إذا كانت هذه النماذج لا يمكن تحويلها لنماذج خطية في المعلمات ، فإنها تسمى نماذج انحدار غير خطية جوهرياً . ومن الآن فصاعداً ، عندما نتحدث عن نماذج انحدار غير خطية ، فإننا نقصد نماذج غير خطية جوهرياً ، وللاختصار سنسمي هذه النماذج NLRM .

لاستيعاب الفرق بين النوعين السابقين من النماذج ، دعنا نسترجع تمرين 6.2 ، 7.2 . في تمرين 6.2 ، النماذج a ، b ، c ، e هي نماذج انحدار خطية لأنها جميعاً نماذج خطية في المعلمات . النموذج d هو نموذج مختلط ، فهو نموذج خطي في β_2 ، وغير خطي في β_1 ، ولكن بوضع $\alpha = \text{Ln}\ \beta_1$ يصبح هذا النموذج خطياً في α ، β_2 .

في تمرين 2.7 ، النموذجان d ، e هما نموذجان غير خطيين جوهرياً ، حيث لا توجد طريقة بسيطة لتحويلهما إلى نماذج خطية . النموذج c هو نموذج انحدار خطي . ماذا عن النموذجين a و b؟

إذا أدخلنا على نموذج a الدالة اللوغاريتمية نحصل على $\text{Ln}\ Y_i = \beta_1 + \beta_2 X_i + u_i$ وهو نموذج خطي في المعلمات . وبالتالي النموذج a هو نموذج انحدار خطي جوهرياً . النموذج b هو مثال لدالة الاحتمال اللوجيستي وسندرسها في الفصل 15 . ونلاحظ أنها تظهر كنموذج انحدار غير خطي ، ولكن بحيلة رياضية بسيطة ستتحول إلى نموذج انحدار خطي ، كالتالي :

$$\ln\left(\frac{1 - Y_i}{Y_i}\right) = \beta_1 + \beta_2 X_i + u_i \qquad (1.1.14)$$

ولهذا فالنموذج b هو نموذج خطي جوهري ، وسنرى أهمية النماذج مثل (1.1.14) في الفصل التالي . ودعنا نستعرض الآن دالة Cobb- Douglas (C-D) الشهيرة للإنتاج ، لنجعل Y = الناتج ، X_2 = العمالة ، X_3 = رأس المال . سنكتب هذه الدالة بثلاث طرق مختلفة :

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i} \qquad (2.1.14)$$

أو

$$\ln Y_i = \alpha + \beta_2 \ln X_{2i} + \beta_2 \ln X_{3i} + u_i \qquad \textbf{(a2.1.14)}$$

بحيث : $\alpha = \text{Ln}\, \beta_1$. في هذا الشكل السابق تكون دالة (C-D) دالة خطية جوهرياً أما الشكل التالي لدالة (C-D) فهو :

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} u_i \qquad \textbf{(3.1.14)}$$

أو

$$\ln Y_i = \alpha + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + \ln u_i \qquad \textbf{(a3.1.14)}$$

بحيث : $\alpha = \text{Ln}\, \beta_1$. هذا النموذج أيضاً يعتبر نموذجاً خطياً في المعلمات . ولكن الآن دعنا نستعرض هذا الشكل لدالة (C-D) :

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} + u_i \qquad \textbf{(4.1.14)}$$

كما لاحظنا يعتبر شكل الدالة (C-D) (a2.1.14) و (a3.1.14) نماذج انحدار خطية جوهرياً (في المعلمات) ، ولكن لا توجد أي طريقة لتحويل (4.1.14) حتى يصبح نموذجاً خطياً في المعلمات(2) . ولهذا فإن (4.1.14) هو نموذج انحدار خطي جوهرياً .

دالة أخرى مشهورة أيضاً ، ولكنها غير خطية جوهرياً هي دالة مرونة الإنتاج البديل الثابتة (CES) والتي تعتبر دالة Cobb- Douglas حالة خاصة منها .

دالة CES تأخذ الشكل التالي :

$$Y_i = A\left[\delta K_i^{-\beta} + (1-\delta) L_i^{-\beta}\right]^{-1/\beta} \qquad \textbf{(5.1.14)}$$

حيث Y = الإنتاج ، K = رأس المال ، L= العمالة ، A = معلمة الثابت ، δ = معلمة التوزيع $(0 < \delta < 1)$ و β = معلمة الاستبدال $(\beta \geq -1)$(3) . بغض النظر عن الطريقة التي ندخل بها المقدار العشوائي للخطأ u_i في دالة الإنتاج ، فإنه لا توجد طريقة تجعلها تمثل نموذج انحدار خطي (في المعلمات) . فهذه الدالة تمثل نموذج انحدار غير خطي جوهرياً .

---

(2) إذا حاولت أخذ التحويلة اللوغاريتمية ، لن تكون صحيحة حيث إن : $\text{Ln}(A+B) \neq \text{Ln}\, A + \text{Ln}\, B.$

(3) لمعرفة خصائص دالة CES ، انظر في :

Michael D. Intriligator, Ronald Bodkin, and Cheng Hsiao, Econometric Models, Techniques, and Applications, 2d ed., Prentice Hall, 1996, pp. 294-295.

2.14 تقدير نماذج الانحدار الخطية وغير الخطية: ESTIMATION OF LINEAR AND NONLINEAR REGRESSION MODELS

لمعرفة الفرق في التقدير بين نماذج الانحدار الخطية وغير الخطية ، دعنا نستعرض النموذجين التاليين :

$$Y_i = \beta_1 + \beta_2 X_i + u_i \quad (1.2.14)$$

$$Y_i = \beta_1 e^{\beta_2 X_i} + u_i \quad (2.2.14)$$

من الواضح أن (1.2.14) هو نموذج انحدار خطي في حين (2.2.14) يمثل نموذج انحدار غير خطي ، انحدار (2.2.14) معروف باسم نموذج الانحدار الأسي ، وأحياناً يستخدم لقياس النمو في متغير ما ، مثل معدل النمو السكاني ، معدل نمو الناتج المحلي ، أو نمو المعروض من المال ، وهكذا .

دعنا نستعرض تقدير معالم النموذجين السابقين بطريقة المربعات الصغرى OLS . في الـ OLS نقوم بتصغير مجموع مربعات الأخطاء (RSS) ، والذي يأخذ الشكل التالي بالنسبة للنموذج (1.2.14) .

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \quad (3.2.14)$$

بحيث إن $\hat{\beta}_1$ ، $\hat{\beta}_2$ هي مقدرات OLS بالنسبة للقيم الحقيقية للـ OLS . بتفاضل المقدار السابق بالنسبة إلى المجهولين الاثنين نحصل على المعادلات الطبيعية السابق ذكرها في (4.1.3) و (5.1.3) وبحل هذه المعادلات آنياً نحصل على مقدرات OLS كما في المعادلة (6.1.3) والمعادلة (7.1.3) . ونلاحظ بشكل عام أن في هذه المعادلات تكون المجاهيل (β's) على الجانب الأيسر والمتغيرات المعلومة (X و Y) على الجانب الأيمن من المعادلة . ونتيجة لهذا نحصل على حلول محددة للمجهولين الاثنين معتمدة على بيانات العينة .

دعنا الآن نستعرض ماذا سيحدث إذا حاولنا تصغير RSS الموجود في (2.2.14) كما موضح في ملحق A14 ، فقرة 1-A14 فإن المعادلات الطبيعية المرتبطة بـ (14.3) ، (15.3) هي كالتالي :

$$\sum Y_i e^{\hat{\beta}_2 X_i} = \beta_1 e^{2\hat{\beta}_2 X_i} \quad (4.2.14)$$

$$\sum Y_i X_i e^{\hat{\beta}_2 X_i} = \hat{\beta}_1 \sum X_i e^{2\hat{\beta}_2 X_i} \quad (5.2.14)$$

وعلى خلاف المعادلات الطبيعية في حالة نموذج الانحدار الخطي ، فإن المعادلات الطبيعية للانحدار غير الخطي تكون فيها المعلمات المجهولة ($\hat{\beta}$'s) على جانبي المعادلة الأيمن والأيسر معاً . بعبارة أخرى ، فإن المجاهيل دوال معرفة في أنفسها وفي البيانات ! وبالتالي على الرغم من إمكانية تطبيق طريقة المربعات الصغرى لتقدير معالم نماذج الانحدار غير الخطية ، فإننا لن نحصل على حلول محددة وواضحة للمجاهيل . وفي حالة تطبيق طريقة المربعات الصغرى لنماذج الانحدار غير الخطية ، فإنها تسمى طريقة المربعات الصغرى غير الخطية (NLLS) والآن . ماهو الحل؟ إجابة ذلك في الفقرات التالية .

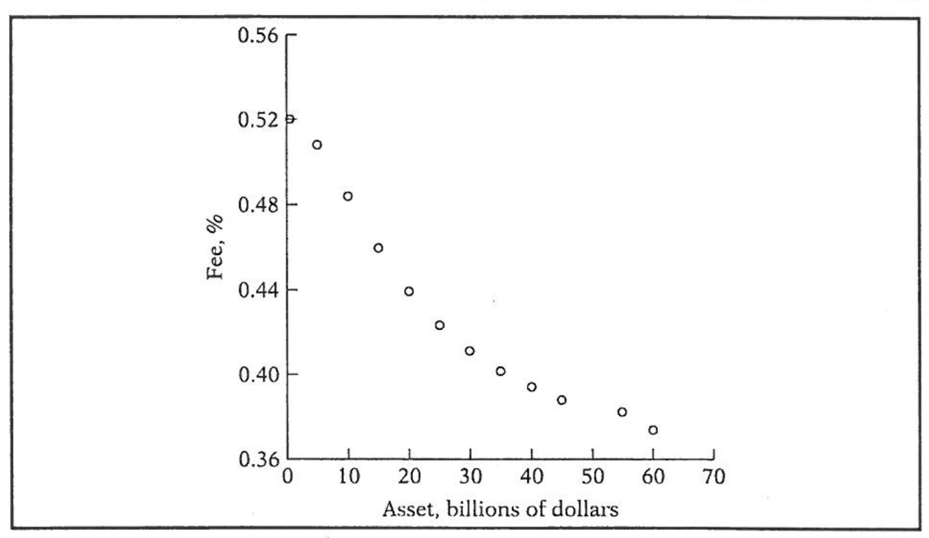
3.14 تقدير نماذج الانحدار غير الخطية: طريقة المحاولة والخطأ
ESTIMATING NONLINEAR REGRESSION MODELS: THE TRIAL- AND- ERROR METHOD

دعنا الآن نفهم الموضوع بشكل أكثر باستخدام مثال واقعي . البيانات الموجودة في جدول (1.14) والخاصة بتكلفة الإدارة التي تدفعها شركة رائدة للتمويل في الولايات المتحدة الأمريكية إلى مستشاريها الاستثماريين الذين يديرون الأصول الصافية للشركة . التكلفة التي تدفعها الشركة تعتمد على القيمة الصافية لأصل التمويل ، وكما ترى كلما زادت القيمة الصافية لأصل التمويل ، كلما قلت تكلفة الاستشارة . وهذا يتضح من الشكل (1.14) لنرى الآن كيف يصلح الانحدار الأسي في (2.2.14) للبيانات المعطاة في جدول (1.14) ، دعنا نجرب طريقة المحاولة والخطأ .

جدول (1.14) التكلفة المدفوعة للاستشارة وحجم الأصل

	Fee, %	Asset*
1	0.520	0.5
2	0.508	5.0
3	0.484	10
4	0.46	15
5	0.4398	20
6	0.4238	25
7	0.4115	30
8	0.402	35
9	0.3944	40
10	0.388	45
11	0.3825	55
12	0.3738	60

*Asset represents net asset value, billions of dollars.



شكل (1.14) العلاقة بين تكلفة الاستشارة وأصل التمويل

دعنا نفترض مبدئياً $\beta_1 = 0.45$ و $\beta_2 = 0.01$. هذه القيم افتراضية ، وقد تعتمد أحياناً على خبرات سابقة ، أو تجارب عملية سابقة ، أو حتى يتم الحصول عليها من نموذج انحدار خطي ، وإن كان لا يصلح لهذه الحالة . ونظراً لمعلومية قيم β_1 و β_2 فإنه يمكن كتابة (2.2.14) كالتالي :

$$u_i = Y_i - \beta_1 e^{\beta_2 X_i} = Y_i - 0.45 e^{0.01 X_i} \qquad \textbf{(1.3.14)}$$

وبالتالي :

$$\sum u_i^2 = \sum (Y_i - 0.45 e^{0.01 X_i})^2 \qquad \textbf{(2.3.14)}$$

وبما أن Y ، X ، β_1 و β_2 كلها قيم معروفة ، نستطيع بسهولة إيجاد مجموع مربعات الأخطاء كما في(2.3.14) (4) . تذكر أنه في الـ OLS يكون هدفنا هو إيجاد هذه القيم غير المعروفة للمعلمات ، بحيث يكون مجموع مربعات الأخطاء أصغر ما يمكن . هذا سيحدث إذا كانت قيم Y المقدرة من النموذج قريبة نوعاً ما من قيم Y الفعلية .

من القيم المعطاة نحصل على $\sum u_i^2 = 0.3044$. ولكن كيف نعرف إذا كان هذا هو فعلاً أقل مجموع مربعات يمكن الحصول عليها للأخطاء؟ ماذا سيحدث إذا تم اختيار قيم أخرى لـ β_1 و β_2 مثلاً 0.5 و 0.01- بالترتيب .

---

(4) لاحظ أن $\sum u_i^2$ تسمى أحياناً بمجموع مربعات الأخطاء وليس مجموع مربعات البواقي ، حيث إن قيم المعلمات مفترض أنها كلها معروفة .

بتكرار نفس الأسلوب السابق ، نحصل على $\sum u_i^2 = 0.0073$ ، وهذا المجموع أقل بشكل واضح مما تم الحصول عليه سابقاً والذي كان يساوي 0.3044 . ولكن كيف نستطيع أن نعرف أننا وصلنا إلى الحد الأدنى الممكن لمجموع مربعات الأخطاء ، مع العلم أن اختيار مجموعة قيم مختلفة β's تجعلنا نحصل على مجموع مربعات خطأ آخر مختلف؟

كما ترى ، فإن طريقة المحاولة والخطأ أو العملية المكررة ممكن عملها بسهولة . وإذا كان الفرد لديه وقت كاف وصبر غير محدود ، فإن طريقة المحاولة والخطأ قد ينتج عنها قيم لـ β_1 ، β_2 تتضمن أقل حد أدنى ممكن لمجموع مربعات الأخطاء . ولكن السؤال الآن : كيف انتقلنا من $(\beta_1 = 0.45, \beta_2 = 0.01)$ إلى $(\beta_1 = 0.5, \beta_2 = -0.1)$؟ واضح أننا نحتاج إلى نوع من نظام الحلول الحسابية والتي توضح كيفية الانتقال من فئة معينة من قيم المجاهيل إلى فئة أخرى قبل أن نتوقف ونحصل على الحل النهائي .

لحسن الحظ ، فإن مثل هذا النظام من الحلول الحسابية متاح ، وسيتم استعراضه في الفقرة التالية .

4.14 أساليب تقدير نماذج الانحدار غير الخطية :
APPROACHES TO ESTIMATING NONLINEAR REGRESSION MODELS

هناك العديد من أساليب وأنظمة الحلول الحسابية الخاصة بـ NLRMs :

1 - البحث المباشر أو المحاولة والخطأ .

2 - الأمثلية المباشرة

3 - الخطية المكررة[5] .

(5) هذه الفقرة تعتمد بشكل كبير على المصادر التالية :
Robert S. Pindyck and Daniel L. Rubinfeld , Econometric Models and Economic Forecasts, 4th ed., McGraw-Hill, 1998, Chap. 10; Norman R. Draper and Harry Smith, Applied Regression Analysis, 3d ed., John Wiley & Sons, 1998, Chap. 24; Arthur S. Goldberger, A Course in econometrics, Harvard University Press, 1991, Chap. 29; Russell Davidson and James MacKinnon, op.cit., pp. 201-207; John Fox, Applied Regression Analysis, Linear Models, and Related Methods, Sage Publications, 1997, pp. 393-400; and Ronald Gallant, Nonlinear Statistical Models, John Wiley and Sons, 1987.

طريقة البحث المباشر أو المحاولة والخطأ أو الطريقة اللاتفاضلية :
Direct Search or Trial-and- Error or Derivative- Free Method

في الفقرة السابقة ، أوضحنا كيف نستخدم هذه الطريقة . وعلى الرغم من سهولتها الواضحة ، حيث إنها لا تتطلب أي استخدام لطرق رياضية ، كما هو الحال في الطرق الأخرى ، فإنه بوجه عام لا تستخدم هذه الطريقة كثيراً . أولاً : إذا كان NLRM يشمل معلمات عديدة فهذه الطريقة تصبح معقدة جداً وصعبة الحساب . فمثلاً إذا كان NLRM يشتمل على 5 معلمات و 25 قيماً بديلة لكل معلمة ، بالتالي فالمطلوب حساب مجموع مربعات أخطاء بعدد $(25)^5 = 9,765,625$ مرة ! . ثانياً : لا يوجد ضمان أن الفئة النهائية لقيم المعلمات سوف تعطي بالضرورة القيم الصغرى المطلقة لمجموع مربعات الأخطاء . وبلغة الحساب ، فإنه قد يمكن الحصول على نهاية صغرى محلية وليست مطلقة . وفي الواقع لا توجد طريقة تضمن نهاية صغرى عامة .

الأمثلية المباشرة : Direct Optimization

في طريقة الأمثلية المباشرة ، نقوم بتفاضل مجموع مربعات الأخطاء بالنسبة إلى كل معامل أو معلمة مجهولة ، ثم نساوي المعادلة التفاضلية بالصفر ، ونقوم بحل المعادلات الطبيعية الناتجة آنياً .

وقد رأينا ذلك من قبل في معادلات (4.2.14) ، (5.2.14) ولكن بالنظر إلى هذه المعادلات ، نرى أنه لا يمكن إيجاد حلول تحليلية واضحة ، ولحل ذلك يتم استخدام بعض الطرق المحددة التي تجري على وتيرة واحدة مكررة . إحدى هذه الطرق يطلق عليها اسم طريقة الهبوط شديد الانحدار . في هذا الكتاب ، لن نقوم بتناول تفاصيل هذه الطريقة ، ولكن القارئ يستطيع إيجاد هذه التفاصيل في المراجع المذكورة . مثل طريقة المحاولة والخطأ ، فإن طريقة الهبوط شديد الانحدار تعتمد أيضاً على اختيار قيم مبدئية للمعالم المجهولة ، ولكنها بعد ذلك تحتوي على خطوات عملية أكثر دقة من طريقة التخمين أو المحاولة والخطأ . أحد عيوب هذه الطريقة ، أنها تؤول إلى القيم النهائية للمعلمات ببطء شديد .

طريقة الخطية المكررة : Iterative Linearization Method

في هذه الطريقة ، نحاول جعل المعادلات غير الخطية خطية حول بعض القيم المبدئية للمعالم ، ثم يتم تقدير هذه المعادلات التي أصبحت خطية بطريقة OLS ، ثم

يتم تعديل القيم المبدئية المختارة . هذه القيم المعدلة تستخدم لإعادة خطية النموذج ، ثم مرة أخرى نقدر المعالم بطريقة OLS ، ثم نعيد تصحيح القيم المقدرة . وتستمر هذه العملية حتى لا يصبح هناك تغير حقيقي في القيم المقدرة في إطار آخر تكرارين . الطريقة الرئيسية التي تستخدم في محاولة خطية المعادلات غير الخطية هي مفكوك متسلسلة تيلور . التفاصيل الرياضية لهذه الطريقة معطاة في ملحق A14 ، فقرة A2.14 . تقرير NLRM باستخدام مفكوك متسلسلة تيلور مبني على عمليتين حسابيتين هما الطريقة المكررة لـ Gauss- Newton والطريقة المكررة لـ Neuton- Raphson . وبما أن واحداً أو أكثر أو كلاً من هاتين الطريقتين أصبحتا متحلتين على حزم الحاسب الآلي ، بالإضافة إلى أن دراسة تفاصيل هاتين الطريقتين ستأخذنا إلى ما هو خارج نطاق هذا الكتاب ، فلن يكون هناك أي داع للخوض في تفاصيلهما[6] .

في الفقرة التالية ، سندرس بعض الأمثلة التي يتم فيها استخدام هذه الطرق .

5.14 أمثلة توضيحية ILLUSTRATIVE EXAMPLES

مثال 1.14

رسوم استشارة التمويل Continued

بالرجوع إلى البيانات المعطاة في جدول (1.14) و NLRM الموجودة في (2.2.14) وباستخدام طريقة Eviews للانحدار غير الخطي ، والتي تعتمد على طريقة الخطية السابق ذكرها[7] . حصلنا على نتائج الانحدار التالية : معاملات الانحدار وأخطاؤها القياسية ، وقيم الـ t الخاصة بها معطى في الجدول التالي .

Variable	Coefficient	Std. error	*t* value	*p* value
Intercept	0.5089	0.0074	68.2246	0.0000
Asset	−0.0059	0.00048	−12.3150	0.0000

$R^2 = 0.9385 \qquad d = 0.3493$

(6) هناك طريقة أخرى تستخدم أحياناً وتسمى طريقة Marquard والتي تعتبر مقارنة بين طريقة الهبوط شديد الانحدار وطريقة الخطية (متسلسلة تيلور) . القارئ المهتم بذلك يمكنه الاستعانة بالمراجع المذكورة للتعرف على تفاصيل هذه الطريقة .

(7) Eviews تعطي ثلاثة خيارات : تسلق التل التربيعي ، Newton- Raphson و Berndthall- hall- haesman . الطريقة التي تستخدم تلقائياً هي تسلق التل التربيعي ، والتي تعتبر أحد أنواع طريقة Newton- Raphson .

من هذه النتائج يمكن كتابة النموذج المقدر كالتالي :

$$\widehat{Fee}_i = 0.5089\, Asset^{-0.0059} \qquad (1.5.14)$$

وقبل أن نناقش هذه النتائج ، يمكننا ملاحظة أنه إذا لم نقم بإعطاء قيم مبدئية للمعلمات حتى تبدأ عملية الخطية فستقوم Eviews بعمل ذلك وتحتاج إلى خمسة تكرارات حتى نحصل على النتائج الموضحة في 1.5.14 .

وعموماً أنت تستطيع إعطاء القيم المبدئية التي تريدها حتى تبدأ العملية ، ولتجربة ذلك ، فقد قمنا باختيار القيم المبدئية $\beta_1 = 0.45$ و $\beta_2 = 0.01$. فحصلنا على نفس النتائج في (1.5.14) ولكن بعد ثمانية تكرارات . ويجب ملاحظة أن عدد التكرارات المطلوب سيكون أقل إذا كانت قيمك المبدئية ليست بعيدة جداً عن القيم النهائية . في بعض الحالات يمكنك اختيار القيم المبدئية عن طريق استخدام انحدار OLS متجاهلاً لعدم الخطية . على سبيل المثال ، باستخدام بيانات جدول (1.14) إذا قمت بانحدار الرسوم على الأصول ، فإن مقدرات OLS بالنسبة لـ β_1 هي 0.5028 وبالنسبة لـ β_2 هي 0.002− وهذه القيم أكثر قرباً من القيم النهائية المعطاة في (1.5.14) (لمزيد من التفاصيل الفنية ، انظر ملحق A14 ، فقرة 3.A14) .

والآن بالنسبة لصفات مقدرات NLLS دعنا نتذكر أنه في حالة نموذج الانحدار الخطي الذي يتبع فيه الخطأ التوزيع الطبيعي كنا قادرين على إيجاد طريقة للاستدلال (مثل : اختيارات الفروض) باستخدام اختيارات t ، F ، X^2 سواء كان حجم العينة صغيراً أو كبيراً . لسوء الحظ هذا ليس الحال في NLRMS حتى ولو كان الخطأ يتبع التوزيع الطبيعي . فمقدرات NLLs ليست لها التوزيع الطبيعي ، وليست غير متحيزة ، وليس لها أقل تباين في العينات ذات الحجم المحدود أو الصغير . وكنتيجة لذلك ، لا تستطيع استخدام اختبار t (لاختبار معنوية إحدى معاملات الانحدار) أو اختبار F (لاختبار المعنوية الكلية لنموذج الانحدار المقدر) وذلك لأننا لا نستطيع الحصول على مقدر غير متحيز لتباين الخطأ σ^2 من البواقي المقدرة . الأكثر من ذلك ، البواقي (الفرق بين قيم Y الحقيقية وقيم Y المقدرة من NLRM) ليس بالضرورة مجموعهما يساوي الصفر RSS, ESS . مجموعهما ليس بالضرورة مساوياً لـ TSS ، وبالتالي $R^2 = \frac{ESS}{TSS}$ قد تكون إحصاء وصفياً ليس له معنى لمثل هذه النماذج ، وعموماً يمكن حساب R^2 كالتالي :

$$R^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum (Y_i - \bar{Y})^2} \qquad (2.5.14)$$

حيث Y هي المتغير المنحدر عليه و $\hat{u}_i = Y_i - \hat{Y}_i$ بحيث إن $\hat{Y}_i$ هي القيم المقدرة لـ Y من الـ NLRM .

وبالتالي الاستدلال الإحصائي عن معلمات الانحدار في الانحدار غير الخطي يعتمد على نظرية العينات الكبيرة . هذه النظرية تنص على أن مقدرات طريقة المربعات الصغرى وطريقة الإمكان الأعظم لنماذج الانحدار غير الخطية والتي يكون فيها الخطأ يتبع التوزيع الطبيعي ويكون حجم العينة كبيراً نسبياً هي مقدرات لها التوزيع الطبيعي تقاربياً وغالباً غير متحيزة . وغالباً لها أقل تباين . ولاحظ أن نظرية العينات الكبيرة تنطبق أيضاً عندما يكون الخطأ لا يتبع التوزيع الطبيعي[8] . باختصار إذن كل عمليات الاستدلال في NLRM تعتمد على العينات الكبيرة أو التقاربية .

بالرجوع إلى مثال 1.14 ، الإحصاء t المعطى في (1.5.14) له معنى فقط إذا تم فهمه في إطار نظرية العينات الكبيرة . وبهذا المعنى نستطيع أن نقول إن تقديرات معاملات الانحدار المعطاة في المعادلة (1.5.14) لها معنوية إحصائية ، ونلاحظ أن العينة في المثال الحالي هي عينة صغيرة الحجم .

بالعودة إلى المعادلة (1.5.14) . كيف ترى معادلة تغيير Y (الرسوم) بالنسبة لـ X (حجم الأصل)؟ باستخدام القواعد الرئيسية في التفاضل يستطيع القارئ أن يستنتج أن معدل تغيير Y بالنسبة لـ X هو :

$$\frac{dY}{dX} = \beta_1\beta_2 e^{\beta_2 X} = (-0.0059)(0.5089)e^{-0.0059X} \qquad (3.5.14)$$

ويتضح أن معدل الزيادة في الرسوم يعتمد على قيم الأصل . فعلى سبيل المثال ، إذا كانت X=20 مليون دولار ، فإن معدل التغيير في الرسوم المتوقع يمكن إيجاده من (2.5.14) ليكون تقريباً 0.0031−% . بالطبع هذه الإجابة تتغير بناء على قيم X المستخدمة في الحساب . للحكم على النموذج باستخدام R^2 المحسوبة من (2.5.14) نجد أن قيم R^2 هي 0.9385 والتي تقترح أن نموذج NLRM المختارة يعبر عن البيانات الموجودة في جدول (1.14) بطريقة جيدة . قيم إحصاء durbin-watson المقدرة هي 0.3493 والتي تعني أن هناك ارتباطًا ذاتياً أو خطأ محتملاً في توصيف النموذج . وعلى الرغم من وجود العديد من الطرق للتغلب على هذه المشكلة كما هو الحال في مشكلة اختلاف التباين في NLRM فإننا لن نتعمق في دراسة هذه المشاكل هنا . وعلى القارئ المهتم بذلك الاستعانة بالمراجع المذكورة .

(8) John Neter, Micheeal H. Kutner, Christopher J. Nachtsheim, and William Wasserman, Applied Regression Analysis, 3d ed., Irwin, 1996, pp. 548-549.

مثال 2.14

دالة إنتاج Cobb-Douglas للاقتصاد المكسيكي :

The Cobb- Douglas production of the Mexican Economy

بالرجوع إلى بيانات تمرين 9.14 نجد أن هذه البيانات خاصة بالاقتصاد المكسيكي في الفترة 1955–1974 . وسوف نرى ما إذا كان NLRM المعطى في (4.1.14) يعبر عن هذه البيانات أم لا .لاحظ أن Y هي الإنتاج ، X_2 العمالة و X_3 رأس المال .

باستخدام Eviews نحصل على نتائج الانحدار التالية بعد 32 عملية تكرارية .

Variable	Coefficient	Std. error	*t* value	*p* value
Intercept	0.5292	0.2712	1.9511	0.0677
Labor	0.1810	0.1412	1.2814	0.2173
Capital	0.8827	0.0708	12.4658	0.0000

$$R^2 = 0.9942 \qquad d = 0.2899$$

وبالتالي فإن دالة Cobb- Douglas هي :

$$\widehat{\text{GDP}}_t = 0.5292\text{Labor}_t^{0.1810}\text{Capital}_t^{0.8827} \qquad (2.5.14)$$

وعند تفسير هذه المعادلة تقاربياً نجد أن معامل انحدار رأس المال هو فقط المعامل المعنوي الوحيد في النموذج . في تمرين 9.14 يكون المطلوب هو مقارنة هذه النتائج مع نتائج أخرى سيتم الحصول عليها من دالة Cobb- Douglas المضاعفة ، كما هو معطى في (2.1.14) .

مثال 3.14

النمو السكاني في الولايات المتحدة ، 1970– 1999

Growth of U.S. population, 1970-1999

بيانات إجمالي السكان للولايات المتحدة في الفترة 1970–1999 معطاة في جدول تمرين 8.14 . نموذج النمو اللوجيستيكي يستخدم عادة في قياس النمو السكاني كالتالي :

$$Y_t = \frac{\beta_1}{1 + e^{(\beta_2+\beta_3 t)}} + u_t \qquad (4.5.14)$$

بحيث إن Y = السكان ، t = الفترة الزمنية (مقاسة وفقاً لترتيب زمني) ، والمعاملات هي β's . ومما يلفت الانتباه في هذا النموذج هو أنه على الرغم من وجود متغيرين اثنين فقط وهما تعداد السكان والزمن ، فإن هناك ثلاثة مجاهيل ، وذلك الأمر

يوضح أن NLRM يمكن أن يحتوي على معلمات أكثر عدداً من المتغيرات .

عينة : 1970–1999

المشاهدات : 30

النتائج بعد عملية تكرارية واحدة فقط معطاة في الجدول التالي :

	Coefficient	Std. error	*t* statistic	*p* value
β_1	1432.738	508.0113	2.8202	0.0089
β_2	1.7986	0.4124	4.3613	0.0002
β_3	−0.0117	0.0008	−14.0658	0.0000

$$R^2 = 0.9997 \quad d = 0.3345$$

وبالتالي النموذج المقدر هو :

$$\hat{Y}_t = \frac{1432.739}{1 + e^{1.7986 - 0.0117t}} \qquad (5.5.14)$$

وبما أن لدينا حجم عينة كبيراً نسبياً ، فإن كل معاملات الانحدار المقدرة معنوية تقاربياً ، والقيمة الصغيرة لإحصاء Durbin-watson يقترح أن مقدار الخطأ محتمل أن يكون مرتبطاً ذاتياً . في تمرين 8.14 يكون المطلوب هو مقارنة النموذج السابق مع النموذج شبه الخطي التالي : $\ln Y_t = \beta_1 + \beta_2 \text{ time} + u_t$ ومطلوب أيضاً حساب معدل نمو السكان باستخدام كل من النموذجين .

6.14 الخلاصة والنتائج: SUMMARY AND CONCLUSIONS

النقاط المهمة في هذا الفصل ، من الممكن تلخيصها في التالي :

1 – على الرغم من سيادة استخدام نموذج الانحدار الخطي في المجال النظري والتطبيقي ، فإنه في بعض الحالات يكون من الضروري استخدام نماذج الانحدار غير الخطي في المعلمات (NLRM) .

2 – الخلفية الرياضية لنماذج الانحدار الخطي واضحة وبسيطة ، بحيث إنها تسمح بالوصول إلى حلول وقيم محددة لمعاملات الانحدار الخاصة بمثل هذه النماذج . نظرية الاستدلال الإحصائي سواء العينات الصغيرة أو كبيرة الحجم لمثل هذه النماذج مبينة بشكل كبير .

3 – على النقيض ، نرى أن في نماذج الانحدار غير الخطية جوهرياً ، فإن المعلمات لا يمكن الحصول عليها بوضوح ، ولابد من تقديرها عددياً من خلال عمليات التكرار .

4 – هناك العديد من الطرق الخاصة بإيجاد مقدرات الـ NLRM مثل : المحاولة والخطأ ، طريقة المربعات الصغرى غير الخطية (NLLS) باستخدام مفكوك متسلسلة تيلور .

5 – حزم البرامج الإلكترونية تحتوي الآن على أكواد خاصة بالانحدار غير الخطي مثل Gauss-Newton ، Newton- Raphson و Marquard . وهي كلها تعتمد على العمليات التكرارية .

6 – مقدرات NLLS ليس لها خصائص الأمثلية في أحجام العينات الصغيرة ، ولكن توجد مثل هذه الخصائص في حالة العينات ذات الأحجام الكبيرة . ولذلك فإن النتائج التي يتم الحصول عليها من الـ NLLS وتكون العينات صغيرة الحجم لابد من تفسيرها بحذر شديد .

7 – مشاكل الارتباط الذاتي ، اختلاف التباين ، وتوصيف النموذج ، قد تتواجد في نماذج الانحدار غير الخطي كما هو الحال في نماذج الانحدار الخطي .

8 – قد قمنا بتوضيح NLLS من خلال أمثلة عديدة . وبالاستعانة بالحزم المختلفة للبرامج الإلكترونية فإن تقديرات NLRM لم تعد شديدة الصعوبة ، كما في السابق . وبالتالي ، على القارئ ألا يتجنب استخدام مثل هذه النماذج عندما تكون هناك ضرورة نظرية أو عملية لذلك .

وفي واقع الأمر ، إذا رجعنا إلى تمرين 10.12 سنجد من المعادلة (1) أنه من المفروض كما ذكرنا تقدير نموذج انحدار غير خطي جوهرياً .

تمـــاريـــن : EXERCISES

أسئلة Questions

1.14 ما هو المقصود بنماذج الانحدار الخطية وغير الخطية جوهرياً؟ اعط بعض الأمثلة .

2.14 إذا كان مقدار الخطأ في دالة إنتاج Cobb- Douglas من الممكن وجوده مضروباً أو مضافاً . كيف تختار بين هذين الاختيارين؟

3.14 ما هو الفرق بين مقدرات OLS وطريقة المربعات الصغرى غير الخطية NLLS؟

4.14 العلاقة بين الضغط ودرجة حرارة البخار المشبع يمكن التعبير عنها كالتالي(9) :

$$Y = \beta_1 (10)^{\beta_2 t/(\gamma + t)} + u_t$$

بحيث إن Y = الضغط ، t = درجة الحرارة . باستخدام طريقة المربعات الصغرى غير الخطية (NLLS) احصل على المعادلات الطبيعية الخاصة بهذا النموذج .

5.14 حدد ما إذا كانت العبارات التالية صحيحة أو خاطئة . فسر إجابتك .

(a) الاستدلال الإحصائي في انحدار NLLS لا يمكن أن يعتمد على اختبارات t ، F ، X^2 العادية حتى إذا كان مقدار الخطأ له التوزيع الطبيعي .

(b) معامل التحديد R^2 ليس له معنى عملي في حالة NLRM .

6.14 كيف يمكن جعل دالة الإنتاج GES المذكورة في هذا الفصل دالة خطية؟ وضح الخطوات الأساسية .

7.14 النماذج التي تصف سلوك المتغير عبر الزمن تسمى عادة نماذج نمو ، هذه النماذج تستخدم في العديد من المجالات مثل علم الاقتصاد ، علم الأحياء ، علم النبات ، علم البيئة ، وعلم السكان . نماذج النمو ممكن أن تأخذ أشكالاً عديدة خطية وغير خطية .

النماذج التالية فيها Y هو المتغير المراد قياس نموه ، و t هو الزمن مقاس بشكل فيه ترتيب زمني ، و u_t هو مقدار الخطأ العشوائي :

a. $Y_t = \beta_1 + \beta_2 t + u_t$

b. $\ln Y_t = \beta_1 + \beta_2 t + u_t$

c. نموذج النمو اللوجيستيكي $Y_t = \frac{\beta_1}{1+\beta_2 e^{-\beta_3 t}} + u_t$

d. compert z نمـــوذج نمـــو $Y_t = \beta_1 e^{-\beta_2 e^{-\beta_3 t}} + u_t$

استعرض خصائص هذه النماذج والمرتبطة بنمو الـ Y وعلاقته بالزمن .

(9) معدلة من : Draper and Smith, op. c.t., p. 554

Problems مسائل :

8.14 بيانات جدول (2.14) خاصة بالتعداد السكاني للولايات المتحدة في الفترة 1970–1999 وهي معطاة بالمليون نسمة . استخدم نماذج النمو المعطاة في تمرين 7.14 وحدد أي هذه النماذج له توفيق أفضل . فسر معلمات النماذج .

جدول (2.14) التعداد السكاني للولايات المتحدة من 78–99

Observation	U.S. population	Time	Observation	U.S. population	Time
1970	205.052	1	1985	238.466	16
1971	207.661	2	1986	240.651	17
1972	209.896	3	1987	242.804	18
1973	211.909	4	1988	245.021	19
1974	213.854	5	1989	247.342	20
1975	215.973	6	1990	249.948	21
1976	218.035	7	1991	252.639	22
1977	220.239	8	1992	255.374	23
1978	222.585	9	1993	258.083	24
1979	225.055	10	1994	260.599	25
1980	227.726	11	1995	263.044	26
1981	229.966	12	1996	265.463	27
1982	232.188	13	1997	268.008	28
1983	234.307	14	1998	270.561	29
1984	236.348	15	1999	273.131	30

Source: Economic Report of the President, 2000.

جدول (3.14) بيانات دوال الإنتاج الخاصة بالاقتصاد المكسيكي

Observation	GDP	Labor	Capital	Observation	GDP	Labor	Capital
1955	114,043	8,310	182,113	1965	212,323	11,746	315,715
1956	120,410	8,529	193,749	1966	226,977	11,521	337,642
1957	129,187	8,738	205,192	1967	241,194	11,540	363,599
1958	134,705	8,952	215,130	1968	260,881	12,066	391,847
1959	139,960	9,171	225,021	1969	277,498	12,297	422,382
1960	150,511	9,569	237,026	1970	296,530	12,955	455,049
1961	157,897	9,527	248,897	1971	306,712	13,338	484,677
1962	165,286	9,662	260,661	1972	329,030	13,738	520,553
1963	178,491	10,334	275,466	1973	354,057	15,924	561,531
1964	199,457	10,981	295,378	1974	374,977	14,154	609,825

المصدر : *Source:* Victor J. Elias, *Sources of Growth: A Study of Seven Latin American Economies*, International Center for Economic Growth, ICS Press, San Francisco, 1992, Tables E-5, E-12, E-14.

ملاحظات : *Notes:* GDP is in millions of 1960 pesos.
Labor is in thousands of people. العمالة معطاة
Capital is in millions of 1960 pesos.

9.14 بيانات جدول (3.14) هي بيانات حقيقية عن GDP ، العمالة ورؤس أموال حقيقية للمكسيك في الفترة من 1955– 1674 . ادرس إمكانية استخدام دالة إنتاج كوب– دوجلاس المضافة والمعطاة في معادلة (4.1.14) . قارن نتائجك مع النتائج التي تم الحصول عليها من استخدام دالة إنتاج كوب– دوجلاس المضروبة والمعطاة في (2.1.14) . أي النتائج أفضل؟

Appendix 14A ملحق A 14

14 1.A استنتاج المعادلات (4.2.14) و (5.2.14)

Derivation of equations (14.2.4) and (14.2.5)

اكتب المعادلة (2.2.14) كالتالي :

$$u_i = Y_i - \beta_1 e^{\beta_2 X_i} \quad (1)$$

وبالتالي

$$\sum u_i^2 = \sum (Y_i - \beta_1 e^{\beta_2 X_i})^2 \quad (2)$$

وبما أن قيم X ، Y معلومة ، فإن مجموع مربعات الأخطاء دالة في β_1 ، β_2 . وبالتالي لتصغير دالة مجموع مربعات الأخطاء ، فإننا نفاضلها جزئياً بالنسبة للمجهولين الاثنين فنحصل على :

$$\frac{\partial \sum u_i^2}{\partial \beta_1} = 2 \sum (Y_i - \beta_1 e^{\beta_2 X_i})(-1 e^{\beta_2 X_i}) \quad (3)$$

$$\frac{\partial \sum u_i^2}{\partial \beta_2} = 2 \sum (Y_i - \beta_1 e^{\beta_2 X_i})(-\beta_1 e^{\beta_2 X_i} \quad (4)$$

باستخدام شرط الأمثلية من الدرجة الأولى لجعل المعادلات السابقة مساوية للصفر ونحلها أنيًا فنحصل على معادلات (4.2.14) و (5.2.14) . لاحظ أنه تفاضل مجموع مربعات الأخطاء معتمداً على قاعدة السلسلة .

14 2.A طريقة الخطية : The Linearization Method

الطلبة الذين درسوا حسابًا من قبل ، يمكنهم استرجاع نظرية تيلور ، والتي تنص على أن أي دالة اختيارية $f(X)$ متصلة ومشتقاتها التفاضلية من الدرجة n متصلة يمكن تقريبها حول النقطة $X= X_0$ بدالة متعددة الحدود ، وللتذكير فإنها تأخذ الشكل التالي :

$$f(X) = \frac{f(X_0)}{0!} + \frac{f'(X_0)(X - X_0)}{1!} + \frac{f''(X_0)(X - X_0)^2}{2!} + \cdots + \frac{f^n(X_0)(X - X_0)^n}{n!} + R$$

بحيث إن $f'(X_0)$ هي المشتقة التفاضلية الأولى للدالة $f(X)$ مسحوبة عند $X= X_0$ و $f''(X_0)$ هي المشتقة التفاضلية الثانية للدالة $f(X)$ مسحوبة عند $X= X_0$ وهكذا ، بحيث إن $n!$ (مضروب n) عبارة عن $n(n-1)(n-2)\ldots 1$ مع الحالة الخاصة $0!= 1$

و R عبارة عن الباقي . عند اعتبار $n=1$ ، نحصل على تقريب خطي ، واختيار $n=2$ نحصل على تقريب متعدد الحدود من الدرجة الثانية . وكما هو متوقع كلما كانت متعددة الحدود من درجة أعلى ، كلما زادت دقة التقدير للدالة الأصلية . المتسلسلة المعطاة في (1) تسمى مفكوك متسلسلة تيلور للدالة $f(X)$ عند النقطة $X=X_0$. وكمثال دعنا نأخذ الدالة التالية :

$$Y = f(X) = \alpha_1 + \alpha_2 X + \alpha_3 X^2 + \alpha_4 X^3$$

بافتراض أننا نريد التقريب عند $X=0$ فنحصل على :

$$f(0) = \alpha_1 \qquad f'(0) = \alpha_2 \qquad f''(0) = 2\alpha_3 \qquad f'''(0) = 6\alpha_4$$

وبالتالي نحصل على التقريبات التالية :

الدرجة الأولى :$Y = \alpha_1 + \frac{f'(0)}{1!} = \alpha_1 + a_2 X + \text{remainder}\,(= \alpha_3 X^2 + \alpha_4 X^3)$

الدرجة الثانية :

$$Y = f(0) + \frac{f'(0)}{1!}X + \frac{f''(0)}{2!}X^2$$
$$= \alpha_1 + \alpha_2 X + \alpha_3 X^2 + \text{remainder}\,(= \alpha_4 X^3)$$

الدرجة الثالثة :

$$Y = \alpha_1 + \alpha_2 X + \alpha_3 X^2 + \alpha_4 X^3$$

نلاحظ أن التقريب من الدرجة الثالثة أعطى المعطاة الأصلية بالضبط . الهدف من التقريب باستخدام متسلسلة تيلور هو اختيار متعددة الحدود ذات الدرجة الأقل ، على أمل أن المقدار الباقي غير تابعي . وغالباً ما يستخدم لتقريب دالة غير خطية بدالة خطية عن طريق إسقاط المقادير الأعلى في الدرجة .

تقريب متسلسلة تيلور من الممكن تعميمه لدوال تحتوي على أكثر من متغير X . على سبيل المثال ، إذا أخذنا الدالة التالية :

$$Y = f(X, Z) \qquad (2)$$

وافترضنا أننا نريد فكها حول $X = 9$ و $Z = b$. نظرية تيلور توضح التالي :

$$f(x,z) = f(a,b) + f_x(a,b)(x-a)$$
$$+ f_z(a,b)\,f(z-b) + \frac{1}{2!}[f_{xx}(a,b)(x-a)^2 \qquad (3)$$
$$- 2 f_{xz}(a,b)(x-a)(z-b) + f_{zz}(a,b)(z-b)^2] + \cdots$$

بحيث إن f_x = التفاضل الجزئي للدالة بالنسبة للـ x (w.r.t) ، f_{xx} = التفاضل الجزئي الثاني للدالة (w.r.t)X ، وبالمثل للمتغير Z ، أما إذا أردنا الحصول على تقريب خطي للدالة سنستخدم أول مقدارين في (3) ، أما إذا أردنا الحصول على تقريب تربيعي أو تقريب من الدرجة الثانية سنستخدم المقادير الثلاثة المذكورة في (3) وهكذا .

14 3.A التقريب الخطي للدالة الأسية المعطاة في (2.2.14)

Linear Approximation of the exponential function given in (14.2.2)

اعتبر الدالة :

$$Y = f(\beta_1, \beta_2) = \beta_1 e^{\beta_2 X} \quad (1)$$

نلاحظ أنه لسهولة التناول تم إسقاط subscript

تذكر أن في هذه الدالة المجاهيل هي معاملات الانحدار β . دعنا نحلل هذه الدالة الخطية في $\beta_1=\beta_1^*$ و $\beta_2=\beta_2^*$. بحيث إن الكميات starred معطاة كقيم ثابتة . لجعل هذه الدالة خطية نقوم بالتالي :

$$Y = f(\beta_1, \beta_2) = f(\beta_1^*, \beta_2^*) + f_{\beta_1}(\beta_1^*, \beta_2^*)(\beta_1 - \beta_1^*) + f_{\beta_2}(\beta_1^*, \beta_2^*)(\beta_2 - \beta_2^*) \quad (2)$$

بحيث إن $f_{\beta 1}$ و $f_{\beta 2}$ هي المشتقات التفاضلية الجزئية للدالة (1) بالنسبة إلى المجاهيل ، وهذه المشتقات سيتم حسابها عن القيم starred (المفترضة) للمعالم المجهولة . لاحظ أننا نستخدم فقط المشتقات التفاضلية من الدرجة الأولى ، حيث إننا نريد جعل هذه الدالة دالة خطية ، والآن افترض أن $\beta_1^* = 0.45, \beta_2^* = 0.01$ والتي ما هي إلا قيم تخمينية للمعاملات الحقيقية فإن :

$$\begin{gathered} f(\beta_1^* = 0.45, \beta_2^* = 0.01) = 0.45e^{0.01X_i} \\ f_{\beta_1} = e^{\beta_2 X_i} \quad \text{and} \quad f_{\beta_2} = \beta_1 X_i e^{\beta_2 X_i} \end{gathered} \quad (3)$$

باستخدام القواعد الرئيسية للتفاضل . نحسب هذه المشتقات التفاضلية عند القيم المعطاة ، ونعوض بذلك في (2) فنحصل على :

$$Y_i = 0.45e^{0.01X_i} + e^{0.01X_i}(\beta_1 - 0.45) + (0.45)X_i e^{0.01X_i}(\beta_2 - 0.01) \quad (4)$$

والتي يمكن كتابتها كالتالي :

$$(Y_i - 0.45e^{0.01X_i}) = e^{0.01X_i}\alpha_1 + 0.45X_i e^{0.01X_i}\alpha_2 \quad (5)$$

بحيث إن :

$$\alpha_1 = (\beta_1 - 0.45) \quad \text{and} \quad \alpha_2 = (\beta_2 - 0.01) \quad (6)$$

والآن دعنا نجـــعـل $Y_i^* = (Y_i - 0.45e^{0.01x_i})$ ، $X_1 = e^{0.01x_i}$ و $X_{2i} = 0.45X_i e^{0.01xi}$ باستـخـدام هذه التـعـريفـات ، وإضـافـة مـقـدار الخطأ u_i يمكن في النـهـاية كـتـابة (5) كالتالي :

$$Y_i^* = \alpha_1 X_{1i} + \alpha_2 X_{2i} + u_i \qquad (7)$$

لدينا الآن نموذج انحدار خطي . وبما أن Y_i^* ، X_{1i} و X_{2i} يمكن حسابها من البيانات ، فإنه يمكن بسهولة تقدير (7) باستخدام OLS ، فنحصل على قيم α_1 ، α_2 . ومن خلال (6) يمكن أن نحصل على :

$$\beta_1 = \hat{\alpha}_1 + 0.45 \qquad \text{and} \qquad \beta_2 = \hat{\alpha}_2 + 0.01 \qquad (8)$$

نحتفظ بقيم β_1^{**} و β_2^{**} بالترتيب ، وتستـخـدم هذه القيم مرة أخرى للبدء في عملية تكرارية كما في (2) ، وبالتالي نحصل على فئة جديدة من قيم المعاملات β . تستطيع الاستمرارية في هذه العمليات المكررة (الخطية) بنفس الطريقة حتى نصل إلى عدم وجود تغير حقيقي في قيم المعاملات β . في مثال 1.14 احتجنا إلى خمس عمليات تكرارية ، ولكن في مثال دالة كوب- دوجلاس المكسيكية احتجنا إلى 32 عمليـة تكرارية . ولكن في النهـاية الخلفيـة المنطقيـة وراء استـخـدام هذه العـمليات التكرارية هو ما تم شرحه سابقاً .

في مثال هيكل رسوم التمويل ، قيم Y^* ، X_1 و X_2 معطاة في (6) موضحة في جدول (4.14) . البيانات الأساسية قيم معطاة في جدول (1.14) . من هذه القيم نتائج الانحدار المرتبطة بـ (7) هي :

Dependent variable: Y*
Method: least squares

Variable	Coefficient	Std. error	t statistic	Probability
X_1	0.039775	0.006229	6.3856	0.0001
X_2	0.001303	0.000157	8.3095	0.0000

$$R^2 = 0.948378 \qquad \text{Durbin–Watson } d = 0.58337$$

الآن باستخدام (8) القارئ يمكنه إثبات أن :

$$\beta_1^{**} = 0.4897 \qquad \text{and} \qquad \beta_2^{**} = 0.0113$$

Y^*	X_1	X_2
0.02249	0.45225	0.22612
0.03238	0.47307	2.3653
0.03158	0.49732	4.9732
0.02964	0.52282	7.8423
0.03043	0.54963	10.9926
0.03439	0.57781	14.4452
0.04109	0.60743	18.2230
0.04965	0.63858	22.3503
0.05923	0.67132	26.8528
0.06918	0.70574	31.7583
0.09402	0.77996	42.8980
0.09939	0.81995	49.1972

قارن هذه النتائج مع التخمين المبدئي 0.45 ، 0.01 للمعلمتين المجهولتين . باستخدام المقدرات الجديدة المعطاة في (9) تستطيع بدء عملية تكرارية مرة أخرى ، وتستمر في ذلك ، حتى يحدث التقريب ، بمعنى أن النتيجة الجديدة للمقدرات لا تختلف كثيراً عما تم الحصول عليه في التكرار السابق . وبالطبع سيكون المطلوب عمليات تكرارية أقل إذا كان التخمين المبدئي أقرب إلى القيم النهائية . ولاحظ أيضاً أننا استخدمنا فقط المقدار الخطي في مفكوك متسلسلة تيلور . إذا كنا نريد استخدام المقدار التربيعي ، أو مقدار من درجة أعلى ، فمن الممكن الوصول إلى القيم النهائية أسرع ، ولكن في العديد من التطبيقات التقريب الخطي أثبت أنه تقريب جيد .

الفصل الخامس عشر

نماذج الانحدار ذات الاستجابة النوعية

QUALITATIVE RESPONSE REGRESSION MODELS

في كل نماذج الانحدار التي قمنا بدراستها حتى الآن، افترضنا أن المتغير التابع أو المتغير المستجيب Y هو متغير كمي، في حين المتغيرات المفسرة هي إما متغيرات كمية أو نوعية (وهمية dummy) أو خليط منهما.

في واقع الأمر، في الفصل التاسع، الذي يتناول المتغيرات الوهمية، رأينا كيف يقدم المتغير المنحدر الوهمي في نموذج الانحدار، والدور الذي يلعبه في مواقف معينة.

وفي هذا الفصل، نتناول عدة نماذج يكون فيها المتغير المنحدر هو متغير نوعي في طبيعته.

وعلى الرغم من زيادة استخدام ذلك في مجالات عديدة من العلوم الاجتماعية والأبحاث الطبية، فإن نماذج الانحدار ذات الاستجابة النوعية تواجه بعض الصعاب في التقدير والتفسير.

وفي هذا الفصل، نتلمس بعض الموضوعات الرئيسية في هذا الشأن، تاركين التفاصيل للكتب الأكثر تخصصاً في ذلك[1].

(1) كدراسة ميدانية للموضوع، قد يجد القارئ المصادر التالية مصادر جيدة:
Statisical Methods for Categorical Data Analysis, Academic Press, 2000 John H. Aldrich and Forrest Nelson, Linear probability, Logit, and Probit Mdels, Sage Publications, 1984; Tim futing Liao, Interpreting Probability Models: Logit, Probit and Other Generalized Linear Models, Sage Publications, 1994, For a very comprehensive review of the literature, see G. S. Maddlals, Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press, 1983.

1.15 طبيعة النماذج ذات الاستجابة النوعية:
THE NATURE OF QUALITATIVE RESPONSE MODELS

دعنا نفترض أننا نريد دراسة قرار المشاركة في قوة العمل (LFP) للرجال البالغين. وحيث إن البالغ إما يعمل أو لا يعمل، فإن (LFP) هي نعم أو لا، وبالتالي المتغير المستجيب أو المنحدر يمكن أن يأخذ فقط قيمتين. فمثلاً، 1 إذا كان الفرد يعمل و 0 إذا كان الفرد لا يعمل. بمعنى أخر فإن المتغير المنحدر هو متغير ثنائي أو مزدوج. الدراسات الاقتصادية عن العمل تقترح أن قرار LFP هو دالة في معدل البطالة، متوسط معدل الأجر، التعليم، دخل الأسرة وهكذا.

وكمثال آخر، انتخابات الرئاسة الأمريكية. دعنا نفترض أن هناك حزبين سياسيين، الديمقراطي والجمهوري. المتغير التابع هنا هو التصويت في الانتخابات لصالح أحد الحزبين. دعنا نضع $Y=1$ إذا كان الفرد يصوت لصالح مرشح الحزب الديمقراطي و $Y=0$ إذا كان التصويت لصالح مرشح الحزب الجمهوري. عدد جيد من الدراسات في هذا المجال قام به الاقتصادي راي فير من جامعة يال (Yale)، إلى جانب دراسات أخرى قام بها العديد من علماء العلوم السياسية[2].

بعض المتغيرات التي تم استخدامها في اختيار الحزب المصوت له هي معدل نمو الـ GDP، معدلات البطالة والتضخم، إذا كان المرشح يدخل الانتخابات لفترة ثانية أو أولى وهكذا. وفي دراستنا الحالية، نجد أن أهم شئ هو أن المتغير المنحدر هو متغير نوعي.

هناك العديد من الأمثلة الأخرى التي يكون فيها المتغير المنحدر متغيراً نوعياً بطبيعته. فمثلاً كون الأسرة تملك أو لا تملك المنزل الذي تقيم به، هل هناك تأمين أو لايوجد تأمين، كل من الزوج والزوجة في قوة العمل أو واحد منهما فقط؟. وبالمثل دراسة هل دواء معين له تأثير فعال في علاج مرض ما أو لا، شركة تجارية تقرر تصفية ماليها من مخزون سلعي أو لا. عضو في مجلس الشيوخ يصوت لصالح نظام ضريبي أو لا، الرئيس الأمريكي يستخدم vetoabill حق الفيتو أو لا. وهكذا.

---

(2) على سبيل المثال
Ray Fair, "Econometrics and Presidential Elections," Joumal of Economic Perspective, Summer 1996, pp. 89-102, and Machael S. Lewis-Beck, Economics and Elections: The Major Westem Democracies, University of Michigan Press, Ann Arbor, 1980.

لا يجب أن نقيد المتغير المستجيب بفئة نعم أو لا، أو نفرض عليه أن يكون فقط متغيراً ثنائيًا. بالعودة إلى مثال انتخابات الرئاسة الأمريكية، افترض أن لدينا ثلاثة أحزاب: ديمقراطي، جمهوري، ومستقل. هنا المتغير المستجيب هو متغير ثلاثي. عموماً، من الممكن أن يكون لدينا المتغير المستجيب له عدة مستويات أو طبقات.

دعنا الآن ندرس مبدئياً المتغير المستجيب الثنائي، ثم نقرر ذلك بناء على النموذج الرئيسي. قبل أن نقوم بذلك، من الضروري معرفة الفرق الأولي بين نموذج الانحدار عندما يكون المتغير المنحدر Y متغيراً كمياً، والنموذج الذي يكون فيه المتغير Y متغيراً نوعياً.

في حالة النموذج الذي يكون فيه المتغير Y متغيراً كميًا، يكون هدفنا هو تقدير القيمة المتوقعة، أو المتوسط، لهذا المتغير بشرط قيم المتغيرات المنحدر عليها. باستخدام مصطلحات الفصل الثاني، ما نريده هو $E(Y_i \mid X_{1i}, X_{2i}, ..., X_{xi})$ بحيث إن X`s هي المتغيرات المنحدر عليها سواء كمية أو نوعية. في النماذج التي يكون فيها y متغيراً نوعياً، يكون هدفنا هو إيجاد احتمال حدوث شيء ما، مثلاً التصويت للمرشح الديمقراطي، امتلاك منزل ما، الانتماء إلى حزب، أو المشاركة في رياضة ما، وهكذا. ولهذا تسمى نماذج الانحدار ذات الاستجابة النوعية بالنماذج الاحتمالية. في المتبقي من هذا الفصل، نحاول الإجابة عن هذه الأسئلة:

1 - كيف نقدر نماذج الانحدار ذات الاستجابة النوعية؟ هل من الممكن تقديرها باستخدام الـ OLS العادية؟

2 - هل توجد مشاكل استدلال؟ بمعنى آخر هل اختيارات الفروض لها طريقة مختلفة عما سبق دراسته حتى الآن؟

3 - إذا كان المتغير المنحدر نوعياً. كيف يمكن قياس جودة التقدير الخاصة بهذه النماذج؟ هل القيمة المحسوبة R^2 لها معنى في مثل هذه النماذج؟

4 - إذا كان لدينا المتغير المنحدر له أكثر من حالتين. كيف نقدر ونفسر نماذج الانحدار ذات الاستجابة المتعددة الطوابق؟ أيضاً كيف نتعامل مع الحالة التي يكون فيها المتغير التابع متغيراً نوعياً ترتيبياً، أي يكون المتغير مرتبًا طبقياً مثل الدراسة المدرسية (أقل من 8 سنوات، 8 إلى 11 سنة، 12 سنة، 13 سنة أو أكثر) أو عندما يكون المتغير التابع نوعياً لا ترتيبيًا، أي يكون متغيراً اسميًا مثل العرق (أسود، أبيض، إسباني، آسيوي أو بخلاف ذلك)؟

5 - كيف نضع نموذجاً للظواهر، مثل عدد الزيارات الطبية لعيادة ما في السنة، عدد براءات الاختراع من شركة ما في سنة معينة، عدد المقالات التي ينشرها أستاذ جامعي في السنة، عدد المكالمات التليفونية التي تصل في فترة 5 دقائق أو عدد السيارات التي تمر في toll booth في فترة 5 دقائق؟ مثل هذه الظواهر تسمى بيانات عددية، أو بيانات نادرة الحدوث، وهي تعتبر أمثلة لعملية (احتمال) بواسون.

في هذا الفصل، نعطي إجابات أساسية على بعض هذه الأسئلة، بعض الموضوعات الأخرى تعتبر متقدمة، وتحتاج إلى خلفية رياضية وإحصائية أكثر مما هو مفترض في هذا الكتاب.

المراجع المذكورة ممكن الرجوع إليها لدراسات أكثر عمقاً.

سنبدأ دراستنا للنماذج ذات الاستجابة النوعية، باعتبار نماذج الانحدار ذات المتغير المستجيب الثنائي. هناك ثلاث طرق ممكنة من خلال تطوير نموذج احتمالي لمتغير مستجيب ثنائي وهي:

1 - النموذج الاحتمالي الخطي (LPM).
2 - نموذج اللوجيت.
3 - نموذج البروبيت.

نظراً للسهولة النسبية وإمكانية تقديرها باستخدام OLS سنذكر أولاً LPM تاركين الطريقتين الأخريين لفقرات أخرى متتالية.

2.15 النموذج الاحتمالي الخطي:
THE LINEAR PROBABILITY MODEL (LPM)

بنفس الفكرة دعنا نعتبر نموذج الانحدار التالي:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad (1.2.15)$$

بحيث إن Y = دخل الأسرة و $Y=1$ إذا كانت الأسرة تمتلك المنزل السكني، أو 0 إذا كانت لا تمتلكه.

نموذج (1.2.15) يبدو كأنه نموذج انحدار خطي تقليدي، ولكن نظراً لأن المتغير المنحدر هو متغير ثنائي أو مزدوج يسمى نموذج احتمال خطي (LPM). ذلك لأن التوقع الشرطي لـ Y_i بمعلومية $E(Y_i \mid X_i), X_i$ يمكن تفسيره بأنه احتمال شرطي للحدث

بمعلومية وقوع Xi وذلك هو $\Pr (Y_i = 1 \mid X_i)$. وبالتالي في المثال الحالي، $E (Y_i \mid X_i)$ يعطي احتمال أن تمتلك الأسرة المنزل السكني، ويكون دخل هذه الأسرة هو الكمية X_i.

تسمية نموذج مثل (1.2.15) بـ LPM يمكن تعليله كالتالي:

نفترض أن $E (u_i)= 0$ (للحصول على مقدرات غير متحيزة) فنحصل كما هو معتاد على:

$$E (Y_i \mid X_i) = \beta_1 + \beta_2 X_i \qquad (2.2.15)$$

الآن إذا كانت P_i هي احتمال أن $Y_i = 1$ (أي أن الحدث يقع) وكانت $1 - P_i$ هي احتمال أن $Y_i = 0$ (أي أن الحدث لا يقع)، فإن المتغير Y له التوزيع الاحتمالي التالي:

Y_i	الاحتمال
0	$1 - P_i$
1	P_i
Total	1

ونقول على المتغير Y_i أنه يتبع توزيع برنولي الاحتمالي.

والآن باستخدام التعريف الرياضي للتوقع، نحصل على:

$$E (Y_i) = 0\,(1 - P_i) + 1(P_i) = P_i \qquad (3.2.15)$$

بمقارنة (2.2.15) مع (3.2.15) نحصل على:

$$E (Y_i \mid X_i) = \beta_1 + \beta_2 X_i = P_i \qquad (4.2.15)$$

بمعنى، أن التوقع الشرطي للنموذج (1.2.15) ما هو في الحقيقة إلا الاحتمال الشرطي لـ Y_i. عموماً توقع المتغير العشوائي الذي له توزيع برنولي هو احتمال أن يساوي هذا المتغير 1. لاحظنا من قبل إذا كان هناك n من المحاولات المستقلة كل منها له احتمال نجاح P واحتمال فشل $1-P$ و X هو عدد مرات النجاح من هذه المحاولات، فإن X يقال إنه يتبع توزيع ذا الحدين. توقع توزيع ذي الحدين هو np وتباينه هو $np(1-P)$. مدلول كلمة توقع النجاح يتم تعريفه وفقاً للمسألة محل الدراسة. بما أن الاحتمال P_i لابد أن يقع ما بين 0 و 1 فإن لدينا القيد التالي:

$$0 \le E (Y_i \mid X_i) \le 1 \qquad (5.2.15)$$

وهذا القيد ينص على أن التوقع الشرطي (الاحتمال الشرطي) لابد أن يقع بين 0 و 1 .

من العرض السابق نرى أن OLS من الممكن أن تمتد بسهولة وتطبق على نماذج الانحدار التي يكون فيها المتغير التابع متغيراً ثنائياً. وبالتالي قد يتوقع أحد أن تطبيق OLS لن يختلف عما سبق، ولكن للأسف هذا ليس صحيحاً، فبالنسبة لـ LPM هناك العديد من المشاكل سنستعرضها كالتالي:

عدم اتباع مقدار الخطأ u_i للتوزيع الطبيعي:

Non- normality of the disturbance u_i

على الرغم من أن OLS لا تفترض أن يكون توزيع الخطأ u_i هو التوزيع الطبيعي، فإننا نفترض ذلك لغرض الاستدلال الإحصائي[3]. ولكن هذا الفرض غير جائز بالنسبة للمقدار u_i في حالة LPMs، حيث إن مقدار الخطأ u_i مثله مثل المتغير Y_i يأخذ قيمتين اثنتين فقط، وبالتالي فإنه أيضاً يتبع توزيع برنولي، وذلك يمكن مشاهدته بسهولة إذا كتبنا (1.2.15) كالتالي:

$$u_i = Y_i - \beta_1 - \beta_2 X_i \qquad (6.2.15)$$

التوزيع الاحتمالي لـ u_i هو:

	u_i	Probability
When $Y_i = 1$	$1 - \beta_1 - \beta_2 X_i$	P_i
When $Y_i = 0$	$-\beta_1 - \beta_2 X_i$	$(1 - P_i)$

(7.2.15)

بالتأكيد u_i لا يمكن افتراض أنه يتبع التوزيع الطبيعي، بل هو يتبع توزيع برنولي. عدم تحقق فرض التوزيع الطبيعي قد لا يكون له أهمية كبيرة، حيث كما نعلم فإن التقدير بنقطة بطريقة OLS سيظل غير متحيز (تذكر إذا كان الهدف هو التقدير بنقطة فإن فرض التوزيع الطبيعي غير ضروري). بالإضافة إلى أنه مع زيادة حجم العينة، فإن النظرية الإحصائية تنص على أن مقدرات OLS تؤول إلى التوزيع الطبيعي بوجه

(3) تذكر أننا أوصينا بضرورة التأكد من فرض التوزيع الطبيعي قبل تطبيقه من خلال الاختيارات المناسبة الخاصة بذلك مثل اختيار جاركو- بيرا.

عام[4]. وكنتيجة لذلك، في العينات الكبيرة الاستدلال الإحصائي لـ LPM سيتبع طريقة OLS العادية المفترضة التوزيع الطبيعي.

اختلاف تباينات الأخطاء: Heteroscedastic variance of the disturbance

من غير الممكن افتراض أن أخطاء LPM ثابتة التباين، حتى إذا كان $E(u_i)=0$ و $Cov(u_i, u_j)=0$ (بمعنى لايوجد ارتباط تسلسلي). وهذا عموماً لا يعتبر شيئاً مفاجئاً. فكما توضح النظرية الإحصائية، بالنسبة لتوزيع برنولي التوقع والتباين النظري هما P و $P(1-P)$ بالترتيب. حيث إن P هي احتمال النجاح (بمعنى حدوث الشيء)، ويكون التباين هو دالة في التوقع. وبالتالي فإن الخطأ يكون مختلف التباين.

بالنسبة لتوزيع مقدار الخطأ في (7.2.15) وبتطبيق تعريف التباين يستطيع القارئ أن يثبت أن (انظر تمرين 10.15)

$$\text{var}(u_i) = P_i(1-P_i) \qquad (8.2.15)$$

بمعنى أن تباين مقدار الخطأ في الـ LPM له خاصية اختلاف التباين، وبما أن $Pi = E(Y_i \mid X_i) = \beta_1 + \beta_2 X_i$. فإن تباين u_i يعتمد على قيم X، وبالتالي فهو ليس ثابت التباين.

نحن نعلم أن ظهور مشكلة اختلاف التباين تجعل مقدرات OLS، على الرغم من أنها تظل غير متحيزة، غير كفء، بمعنى أنه لا تعد لها أقل تباين. ولكن مشكلة اختلاف التباين، مثل مشكلة فرض التوزيع الطبيعي، ليست مشكلة لا يمكن تخطيها. في الفصل الحادي عشر، ناقشنا عدداً من الطرق التي يمكن بها معالجة مشكلة اختلاف التباين، وبما أن تباين u_i يعتمد على $E(Y_i \mid X_i)$ أحد الطرق التي يمكن بها معالجة مشكلة اختلاف التباين هي تحويل النموذج (1.2.15) بالقسمة على:

$$\sqrt{E(Y_i/X_i)[1 - E(Y_i/X_i)]} = \sqrt{P_i(1 - P_i)} = \text{say} \sqrt{w_i}$$

---

(4) الإثبات معتمد على نظرية النزعة المركزية والرد موجود في:
E.Malinvand, Statistical Methods of Economics, Rand McNally, Chicago, 1966, pp. 195-197.
إذا كانت المتغيرات المنحدرة هي متغيرات عشوائية ولها التوزيع الطبيعي، فإن اختيارات F، t ممكن استخدامها حتى لو كان الخطأ لا يتبع التوزيع الطبيعي. ويجب أن تضع في الاعتبار كلما زاد حجم العينة كلما يؤول توزيع ذي الحدين إلى التوزيع الطبيعي.

وبالتالي :

$$\frac{Y_i}{\sqrt{w_i}} = \frac{\beta_1}{\sqrt{w_i}} + \beta_2 \frac{X_i}{\sqrt{w_i}} + \frac{u_i}{\sqrt{w_i}} \qquad (9.2.15)$$

وبسهولة يمكن إثبات أن مقدار الخطأ بعد التحويل في (9.2.15) هو ثابت التباين، وبالتالي بعد تقدير (1.2.15) من الممكن تقدير (9.2.15) باستخدام OLS والتي ما هي إلا مقدرات (الوزن) المربعات الصغرى المرجحة (WLS) باعتبار W_i هو الترجيح في الجانب النظري ما تم عرضه حتى الآن من مشاكل مازال من الممكن التعامل معه. ولكن في الجانب التطبيقي المقدار الحقيقي $E(Y_i | X_i)$ غير معلوم، وبالتالي الأوزان W_i غير معلومة أيضاً ولتقدير W_i تستطيع استخدام الطريقة التالية، والتي تستعمل على خطوتين رئيستين وهما[5] :

الخطوة الأولى :

استخدم انحدار (OLS) كما في (1.2.15) بغض النظر عن مشكلة اختلاف التباين، واحصل على $\hat{Y}_i$ كمقدار للمقدار الحقيقي $E(Y_i | X_i)$ ثم احصل على $\hat{W}_i = \hat{Y}_i(1 - \hat{Y}_i)$ واعتبرها تقديراً لـ W_i.

الخطوة الثانية :

استخدم مقدرات W_i لتحويل البيانات كما هو موضح في (9.2.15) وقدر المعادلة المحولة باستخدام OLS (المربعات الصغرى المرجحة).

سنقوم سريعاً بشرح هذه الطريقة على المثال التالي، ولكن هناك مشكلة أخرى نريد استعراضها أولاً ومتعلقة بـ LPM.

عدم تحقق شرط $0 \le E(Y_i | X) \le 1$: Nonfulfillment of $0 \le E(Y_i | X) \le 1$

بما أن $E(Y_i | X)$ في نماذج الاحتمال الخطية تقيس الاحتمال الشرطي للحدث y بمعلومية X، فمن الضروري أن يقع هذا التوقع بين 0، 1. على الرغم من بديهية هذا الشرط، فإنه لا يوجد ما يضمن أن $\hat{Y}_i$ والذي يغير تقدير $E(Y_i | X_i)$ سيحقق بالضرورة

(5) لمعرفة أكثر عمقاً لهذه الطريقة، انظر في المرجع التالي:
S.Goldberger, Econometric Theory, John Wiley & Sons, New York, 1964, pp. 249-250.
التغييرات الخاصة بهذه الطريقة تعتمد على العينات الكبيرة التي تم مناقشتها من قبل تحت عنوان تقديرات طريقة المربعات الصغرى العامة في الفصل الخاص باختلاف التباين (انظر فقرة 6.11).

هذا الشرط وهذه مشكلة واقعية خاصة بتقديرات OLS الخاصة بـ LPM. هناك طريقتان لمعرفة ما إذا كان التقدير $\hat{Y}_i$ سيقع بين 0، 1 أو لا. إحدى هذه الطرق هي تقدير الـ LPM باستخدام طريقة OLS العادية ومعرفة ما إذا كان التقدير Y_i يقع ما بين 0، 1 أو لا. إذا كانت هناك قيم أقل من 0 (قيم سالبة) نعتبر $\hat{Y}_i = 0$ في مثل هذا الحالات، إما إذا كانت القيم أكبر من 1 فيمكن اعتبار هذه القيم مساوية لـ 1. الطريقة الثانية هي استخدام أسلوب رياضي يضمن أن تكون تقديرات الاحتمالات الشرطية $\hat{Y}_i$ تقع ما بين 0، 1. نماذج اللوجيت والبروبيت التي سيتم مناقشتها لاحقاً. في هذا الفصل ستضمن أن تكون الاحتمالات المقدرة بالضرورة ما بين الحدود المنطقية المطلوبة 0، 1.

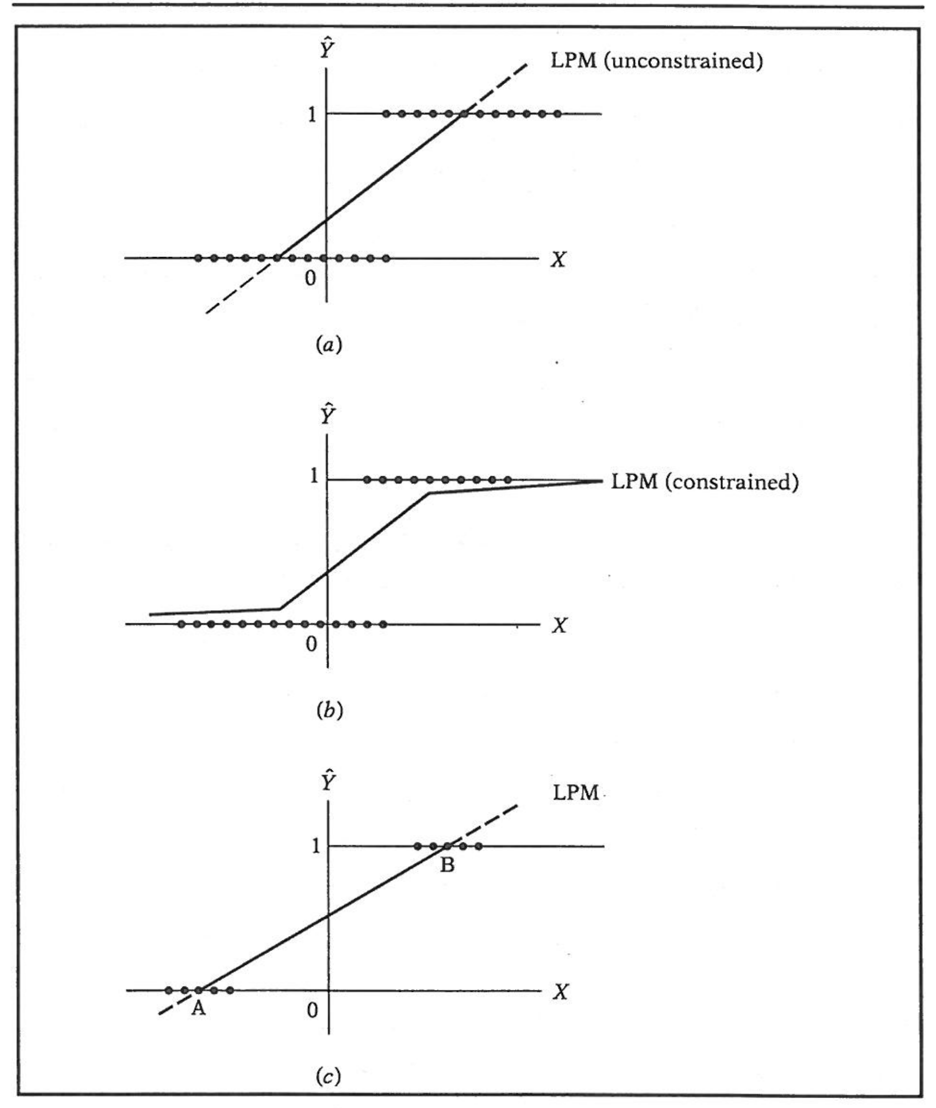
الجدل حول قيمة R^2 كمقياس لجودة التوفيق:

Questionable value of R^2 as a measure of goodness of fit

أهمية القيمة المألوفة R^2 محدودة في إطار النماذج ذات الاستجابة النوعية المزدوجة. ولمعرفة أسباب ذلك، دعنا ندرس الشكل التالي بالنسبة إلى قيمة معينة للـ X، فإن Y إما 0 أو 1. وبالتالي كل قيم الـ Y إما ستقع على المحور السيني أو ستقع على الخط المرتبط بـ 1. وبالتالي بوجه عام لا يوجد *LPM* متوقع أن يصف هذا الشكل جيداً. سواء كان *LPM* غير المقيد (شكل a1.15) أما *LPM* المقيد أو المبتور (شكل b1.15)، تقدير PM يتم بشكل ما يضمن عدم الخروج عن الحدود المنطقية 1−0. وكنتيجة لذلك، القيمة المألوفة R^2 من المتوقع أن تكون أقل بكثير من 1 لمثل هذه النماذج. في العديد من التطبيقات العملية تكون R^2 ما بين 0.2 و 0.6، R^2 لمثل هذه النماذج ستكون كبيرة مثلاً تزيد عن 0.8 في حالة ما إذا كان الشكل الأصلي قريباً جداً حول النقاط A و B (شكل c1.15) وفي مثل هذه الحالات، من السهل تبديل الخط المستقيم بالوصل ما بين النقاط A، B، وتكون القيمة المقدرة لـ Y_i قريبة جداً إما من 0 أو 1.

مثل هذه الأسباب جعلت چون ألدريش وفوريست نيلسون يقولان إن "استخدام معامل التحديد كإحصاء تلخيصي لابد أن يتم تجنبه في حالة النماذج التي يكون فيها المتغير التابع متغيراً نوعيًا[(6)]".

(6) Aldrich and Nelson, op. cit., p15 لمعرفة المزيد عن مقاييس أخرى لجودة التقدير عندما يكون المتغير المنحدر متغيراً وهمياً اقرأ في "Qualitative Response Models," Journal of Economic Literature, vol. 19, 1981, pp. 331-354.



شكل (1.15) نماذج الاحتمالات الخطية

LPM: مثال رقمي:

LPM: A Numerical example

لشرح بعض النقاط الخاصة بـ *LPM* والمذكورة في الفقرة السابقة، دعنا نستعرض مثالاً رقميًا. جدول (1.15) يعطي بيانات عن ملكية المنزل السكني حيث *Y* (=1 امتلاك المنزل، 0 = عدم امتلاك المنزل) و X هي دخل الأسرة (مقاس بالألف دولار) لـ 40 أسرة. من البيانات يكون تقدير *LPM* باستخدام *OLS* هو كالتالي:

$$\hat{Y}_i = -0.9457 + 0.1021X_i$$
$$(0.1228) \quad (0.0082) \qquad (10.2.15)$$
$$t = (-7.6984) \quad (12.515) \quad R^2 = 0.8048$$

جدول (1.15)

بيانات افتراضية عن ملكية المنزل السكني (Y=1 تمثل امتلاك المنزل ، 0 بخلاف ذلك) والدخل X (مقاس بالألف دولار) .

Family	Y	X	Family	Y	X
1	0	8	21	1	22
2	1	16	22	1	16
3	1	18	23	0	12
4	0	11	24	0	11
5	0	12	25	1	16
6	1	19	26	0	11
7	1	20	27	1	20
8	0	13	28	1	18
9	0	9	29	0	11
10	0	10	30	0	10
11	1	17	31	1	17
12	1	18	32	0	13
13	0	14	33	1	21
14	1	20	34	1	20
15	0	6	35	0	11
16	1	19	36	0	8
17	1	16	37	1	17
18	0	10	38	1	16
19	0	8	39	0	7
20	1	18	40	1	17

أولاً-دعنا نفسر هذا الانحدار: الجزء المقطوع من المحور الصادي 0.9457 - يعطي احتمال الأسرة معدومة الدخل (دخل= 0) أن تمتلك المنزل السكني وبما أن هذه القيمة سالبة، وبما أن الاحتمال لا يمكن أن يكون سالباً نتعامل مع هذه القيمة وكأنها 0 وهو ما يعتبر منطقياً في المثال الحالي[7] .

قيمة الميل 0.1021 تعني أنه لكل وحدة واحدة من التغير في الدخل (هنا 1000 دولار) تزيد في المتوسط احتمال امتلاك المنزل السكني بـ 0.1021 أو حوالي عشرة في المائة (10%) . بالطبع يمكن تقدير احتمال امتلاك المنزل السكني بمعلومية قيمة معينة من الدخل من المعادلة (10.2.15) وبالتالي لـ X=12($12.000) فإن تقدير احتمال امتلاك المنزل هو:

$$(\hat{Y}_i|X = 12) = -0.9457 + 12(0.1021) = 0.2795$$

(7) من الممكن تفسير القيمة السالبة على أنها عدم احتمال امتلاك المنزل السكني طالما الدخل يساوي 0.

ومعنى ذلك أن احتمال امتلاك الأسرة للمنزل السكني ويكون دخلها (12.000$) هو 28% . جدول (2.15) يعطي تقديرات الاحتمال، $\hat{Y}_i$ ، بقيم مختلفة من الدخل موضحة في الجدول. الشيء الملحوظ في هذا الجدول أن هناك ست قيم مقدرة سالبة، وست قيم أخرى أكبر من 1 . وذلك يؤكد على ما سبق ذكره من أنه على الرغم من أن $E(Y_i|X)$ قيمة موجبة أقل من الواحد، فإن تقديرات $\hat{Y}_i$ ليست بالضرورة موجبة أو أقل من 1 . وهذا أحد أسباب عدم التوصية باستخدام *LPM* عندما يكون المتغير التابع متغيراً ثنائياً.

حتى إذا كانت القيم المقدرة Y_i كلها قيم موجبة وأقل من 1 . الـ LPM تعاني من مشكلة اختلاف التباين، والتي يمكن رؤيتها في (8.2.15) . ونتيجة لذلك لا يمكن الاعتماد على التقديرات القياسية للأخطاء والمذكورة في (10.12.15) . (ما السبب في ذلك؟) ولكن من الممكن استخدام طريقة المربعات الصغرى المرجحة (WLS) المذكورة سابقاً للحصول على مقدرات أكثر كفاءة للأخطاء القياسية الأوزان، $\hat{W}_i$ المطلوبة في تطبيق لـ WLS معطاة في جدول (2.15) ولكن لاحظ أنه بما أن بعض قيم Y_i قيم سالبة والبعض الآخر أكبر من 1 ، التقديرات $\hat{W}_i$ الخاصة بهذه القيم ستكون سالبة، وبالتالي لا تستطيع استخدام هذه المشاهدات في WLS (لماذا؟) وكنتيجة لذلك نقلل عدد المشاهدات من 40 إلى 28 مشاهدة في مثالنا الحالي[8].

جدول (2.15) قيم Y الحقيقية ، Y المقدرة ، الأوزان W_i لمثال امتلاك المنزل السكني

Y_i	$\hat{Y}_i$	$\hat{w}_i$‡	$\sqrt{\hat{w}_i}$	Y_i	$\hat{Y}_i$	$\hat{w}_i$‡	$\sqrt{\hat{w}_i}$
0	−0.129*			1	1.301†		
1	0.688	0.2146	0.4633	1	0.688	0.2147	0.4633
1	0.893	0.0956	0.3091	0	0.280	0.2016	0.4990
0	0.178	0.1463	0.3825	0	0.178	0.1463	0.3825
0	0.280	0.2016	0.4490	1	0.688	0.2147	0.4633
1	0.995	0.00498	0.0705	0	0.178	0.1463	0.3825
1	1.098†			1	1.097†		
0	0.382	0.2361	0.4859	1	0.893	0.0956	0.3091
0	−0.0265*			0	0.178	0.1463	0.3825
0	0.076	0.0702	0.2650	0	0.076	0.0702	0.2650
1	0.791	0.1653	0.4066	1	0.791	0.1653	0.4055
1	0.893	0.0956	0.3091	0	0.382	0.2361	0.4859
0	0.484	0.2497	0.4997	1	1.199†		
1	1.097†			1	1.097†		
0	−0.333*			0	0.178	0.1463	0.3825
1	0.995	0.00498	0.0705	0	−0.129*		
1	0.688	0.2147	0.4633	1	0.791	0.1653	0.4066
0	0.076	0.0702	0.2650	1	0.688	0.2147	0.4633
0	−0.129*			0	−0.231*		
1	0.893	0.0956	0.3091	1	0.791	0.1653	0.4066

* Treated as zero to avoid probabilities being negative.
† Treated as unity to avoid probabilities exceeding one.
‡ $\hat{Y}_i(1-\hat{Y}_i)$.

(8) لتجنب مشكلة تحليل درجات الحرية، ممكن اعتبار $\hat{Y}_i = 0.01$ عندما تكون القيمة المقدرة لـ Y_i سالبة و$\hat{Y}_i = 0.99$ عندما تكون أكبر من 1 . انظر تمرين 1.15 .

بعد حذف هذه المشاهدات سيكون معادلة انحدار WLS هي :

$$\frac{\hat{Y}_i}{\sqrt{\hat{w}_i}} = -1.2456\frac{1}{\sqrt{\hat{w}_i}} + 0.1196\frac{X_i}{\sqrt{\hat{w}_i}} \qquad (11.2.15)$$

$$(0.1206) \qquad (0.0069)$$

$$t = (-10.332) \qquad (17.454) \qquad R^2 = 0.9214$$

بمقارنة هذه النتائج مع (10.12.15)، نجد أن t = تقديرات الأخطاء القياسية أصغر، وبالتالي القيمة t (القيمة المطلقة) أكبر، ولكن مع الوضع في الاعتبار أنه حتى نصل لهذه النتائج قمنا باستبعاد 12 مشاهدة. وأيضاً بما أن قيمة W_i هي قيمة مقدرة فإن اختبارات الفروض صحيحة فقط في حالة العينات الكبيرة (انظر الفصل الحادي عشر).

3.15 تطبيقات على LPM : A PPLICATIONS OF LPM

قبل إمكانية التقدير لنماذج اللوجيت والبروبيت (سيتم مناقشتها لاحقاً) باستخدام حزم البرامج الإلكترونية، كانت LPM كثيرة التطبيق نظراً لسهولتها النسبية. دعنا نستعرض الآن بعض هذه التطبيقات :

مثال 1.15

دراسة كوهين- ريا- ليرمان[9] Cohen- Rea- Lerman study

في دراسة تم إعدادها لوزارة العمل بالولايات المتحدة الأمريكية، قام كل من كوهين، ريا، وليرمان بدراسة المشاركة في قوة العمل داخل عدد من الأعمال المختلفة كدالة في عدد من العوامل الاقتصادية والاجتماعية والسكانية. في كل الانحدارات التي قاموا بها، تم اعتبار المتغير التابع متغيراً وهمياً يأخذ القيمة 1 إذا كان الشخص داخل قوة العمل، و 0 إذا كان غير موجود في قوة العمل. في جدول (3.15) نجد النتائج الخاصة بمعادلة انحدار واحدة من الانحدارات العديدة التي قاموا بها وكان فيها المتغير التابع متغيراً وهمياً.

قبل التعليق على النتائج، دعنا نلاحظ التالي: الانحدار الحالي تم تقديره باستخدام الـ OLS. لتصحيح اختلاف التباين، استخدموا الطريقة ذات الخطوات السابق ذكرها في بعض الانحدارات. ولكن وجدوا أن الأخطاء القياسية الخاصة بالمقدرات لم تختلف بشكل حقيقي عن نظيرها بدون تصحيح اختلاف التباين. ربما تكون هذه النتيجة بسبب حجم العينة الكبير والذي تقريباً يساوي 25,000. القيمة

(9) مرجع Malcolm S. Cohen, samuel A. Rea, Jr., and Robert I. Lerman, A Mico Model of Labor Supply, Bls Staff Paper 4. U.S. Department of Labor, 1970.

المقدرة لـ t من الممكن استخدامها في اختبار المعنوية بطريقة الـ OLS العادية حتى ولو كان مقدار الخطأ يأخذ قيماً ثنائية. القيمة المقدرة لـ R^2 وهي 0.175 قد تبدو قليلة نوعاً ما، ولكن بالنظر إلى حجم العينة الكبير، تظل هذه القيمة لـ R^2 معنوية على أساس اختيار F المعطى في الفقرة 5.8. وفي النهاية لاحظ كيف خلط الباحثون بين قيم متغيرات كمية وأخرى نوعية، وكيف أخذوا في الاعتبار التأثير الناتج من التفاعل بين هذه المتغيرات.

بالعودة إلى التعليق وتفسير النتائج، نجد أن كل معاملات الميل تعطي معدل التغير في الاحتمال الشرطي لظهور الحدث وفقاً لتغير قيم المتغير المفسر بوحدة واحدة. على سبيل المثال، المعامل 0.2753− والمرتبط بالمتغير "السن 65 سنة وأكثر" يعني بافتراض ثبات باقي العوامل، فإن احتمال المشاركة في قوة العمل بالنسبة للسيدات في هذه الفئة العمرية أقل بحوالي 27% (بالمقارنة مع الفئة الأساسية للسيدات بين 22 إلى 54 سنة). بنفس الطريقة، معامل 0.3061 والمرتبط بالمتغير "الدراسة التعليمية لـ 16 سنة فأكثر" يعني بافتراض ثبات باقي العوامل، فإن احتمال السيدات يمثل هذا المستوى التعليمي أن يشاركوا في قوة العمل أكثر بحوالي 31% (بالمقارنة مع السيدات اللاتي لهن مستوى تعليمي أقل من 5 سنوات تعليمية، الفئة الأساسية).

والآن دعنا ندرس مقدار التفاعل بين السن والحالة الاجتماعية. الجدول يوضح أن احتمالات المشاركة في القوة التعليمية أعلى بحوالي 29% للسيدات اللاتي لم يتزوجن نهائياً (مقارنة بالفئة الأساسية) وأقل بحوالي 28% للسيدات اللاتي تكون أعمارهن 65 سنة فأكثر (مرة أخرى مقارنة بالفئة الأساسية) ولكن احتمال المشاركة في قوة العمل من السيدات اللاتي لم يتزوجن من قبل مطلقاً وعمرهن 65 سنة فأكثر أقل بحوالي 20% بالمقارنة مع الفئة الأساسية. وهذا يعني أن السيدات اللاتي أعمارهن 65 سنة فأكثر ولم يتزوجن من قبل من المحتمل أكثر مشاركتهن في قوة العمل عن هؤلاء لكن اللاتي أعمارهن 65سنة فأكثر ومتزوجات (ينتمين إلى الفئة الأخرى) وباتباع نفس الطريقة يمكن للقارئ بسهولة تفسير باقي المعاملات المعطاة في جدول (3.15). من البيانات المتاحة، من السهل الحصول على مقدرات للاحتمالات الشرطية للمشاركة في قوة العمل وفقاً للفئات المختلفة. وبالتالي إذا أردنا إيجاد احتمال المشاركة في قوة العمل بالنسبة لسيدة متزوجة، العمر بين 22 و 54، وسنوات تعليم ما بين 12 إلى 15 سنة، معدل بطالة ما بين 2.5% إلى 3.4%، تغير وظيفي بمعدل 3.59% إلى 6.49%، احتمالات فرص بديلة للعمل بحوالي 74% فأكثر Filow بحوالي 7500$ فأكثر نحصل على:

$$0.4368 + 0.1523 + 0.2231 - 0.0213 + 0.0571 - 0.2455 = 0.6326$$

بمعنى آخر، احتمال المشاركة في قوة العمل بالنسبة للسيدات بالمواصفات السابقة مقدر بحوالي 63%.

جدول (3.15) المشاركة في قوة العمل

انحدار السيدات ، سن 22 فأكثر ، يسكن في

(المتغير التابع : التواجد أو عدم التواجد في قوة العمل في سنة 1966) .

Explanatory variable		Coefficient	*t* ratio
Constant		0.4368	15.4
Marital status			
Married, spouse present		—	—
Married, other		0.1523	13.8
Never married		0.2915	22.0
Age			
22–54		—	—
55–64		−0.0594	−5.7
65 and over		−0.2753	−9.0
Years of schooling			
0–4		—	—
5–8		0.1255	5.8
9–11		0.1704	7.9
12–15		0.2231	10.6
16 and over		0.3061	13.3
Unemployment rate (1966), %			
Under 2.5		—	—
2.5–3.4		−0.0213	−1.6
3.5–4.0		−0.0269	−2.0
4.1–5.0		−0.0291	−2.2
5.1 and over		−0.0311	−2.4
Employment change (1965–1966), %			
Under 3.5		—	—
3.5–6.49		0.0301	3.2
6.5 and over		0.0529	5.1
Relative employment opportunities, %			
Under 62		—	—
62–73.9		0.0381	3.2
74 and over		0.0571	3.2
FILOW, $			
Less than 1,500 and negative		—	—
1,500–7,499		−0.1451	−15.4
7,500 and over		−0.2455	−24.4
Interaction (marital status and age)			
Marital status	Age		
Other	55–64	−0.0406	−2.1
Other	65 and over	−0.1391	−7.4
Never married	55–64	−0.1104	−3.3
Never married	65 and over	−0.2045	−6.4
Interaction (age and years of schooling completed)			
Age	Years of schooling		
65 and over	5–8	−0.0885	−2.8
65 and over	9–11	−0.0848	−2.4
65 and over	12–15	−0.1288	−4.0
65 and over	16 and over	−0.1628	−3.6
	$R^2 = 0.175$		
No. of observations = 25,153			

لاحظ أن: Filow الدخل الأسري less own دخل الأجور ورواتب تعبر عن الفئة الأساسية أو المحذوفة

المصدر: Malcolm S. Cohen, Samuel A. Rea, Jr., and Robert I. Lerman, A Micro Model of Labor supply, BLS Staff Paper 4, U.S. Department of Labor, 1970, Table F-6. PP. 212-213.

مثال 2.15

التنبؤ بتصنيف السند المالي Predicting a bond rating

قام جوزيف كابليري بتقدير نموذج للتنبؤ بتصنيف السند المالي معتمداً على بيانات مجمعة من سلسلة زمنية، وبيانات جدولية خاصة بـ 200 سند مالي، سواء كان عالي أو متوسط الجودة خلال الفترة الزمنية 1961–1966 . هذا النموذج يأخذ الشكل التالي(10) :

$$Y_i = \beta_1 + \beta_2 X_{2i}^2 + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_i$$

بحيث إن :

$Y_i = 1$ إذا كان السند عالي الجودة Aa (تصنيف مودي)

0 = إذا كان السند متوسط الجودة Baa (تصنيف مودي)

$X_2 =$ معدل ديون رأس المال، كمقياس للنفوذ

$$= \frac{\text{القيم الدولارية لدين طويل المدى}}{\text{القيمة الدولارية لإجمالي رأس المال}} \times 100$$

$X_3 =$ معدل الربح

$$= \frac{\text{القيم الدولارية للدخل بعد دفع الضرائب}}{\text{القيمة الدولارية لإجمالي صافي الأصول}} \times 100$$

$X_4 =$ الانحراف المعياري لمعدل الربح، كمقياس للتباين في معدل الربح

$X_5 =$ إجمالي صافي الأصول (الألف دولار) كمقياس للحجم

هناك توقع لأن تكون قيم β_2 و β_4 قيماً سالبة (لماذا؟) ويتوقع أن تكون β_3 و β_5 قيماً موجبة.

بعد تعديل اختلاف التباين والارتباط الذاتي من الدرجة الأولى، حصل كابليري على النتائج التالية(11) :

$$\hat{Y}_i = 0.6860 - 0.0179 X_{2i}^2 + 0.0486 X_{3i} + 0.0572 X_{4i} + 0.378(E\text{-}7) X_5$$

$$(0.1775) \quad (0.0024) \quad (0.0486) \quad (0.0178) \quad (0.039)(E\text{-}8) \qquad (1.3.15)$$

$$R^2 = 0.6933$$

لاحظ أن : 0.378E تعني 0.000378 وهكذا كل المعاملات باستثناء X_4 لها الإشارة الصحيحة. ومتروك لطلبة الاستثمار التعليل لماذا يأخذ معامل تباين معدل الربح قيمة

(10) Joseph Cappelleri, "Predictiong a Bond Rating," unpublished term paper, C.U.N.Y. The model used in the paper is a modification of the model used by Thomas F. Pogue and Robert M. Soldofsky, "What Is in a Bond Rating?" Journal of Financial and Quantitative Analysis, June 1969, pp. 201-228.

(11) بعض قيم الاحتمالات المقدرة قبل التعامل مع مشكلة اختلاف التباين كانت قيماً سالبة وبعضها كان أكبر من 1 في مثل هذه الحالات كان مفترضاً لها قيم مابين 0.01 و 0.99 بالترتيب لتسهيل حساب الأوزان.

موجبة، من الممكن أن يتوقع الفرد أنه كلما زاد التباين في الربح كلما قل احتمال أن يعطي تصنيفاً عالي الجودة Aa، وذلك بافتراض ثبات العوامل الأولى.

تفسير معادلة الانحدار يتم بشكل تقليدي. على سبيل المثال، 0.0486 والمرتبطة بالمتغير X_3 تعني أنه بافتراض ثبات العوامل الأخرى فإن كل زيادة بوحدة واحدة في معدل الربح ستؤدي في المتوسط إلى زيادة بحوالي 0.05 في احتمال أن يأخذ السند المالي التصنيف عالي الجودة Aa. وبالمثل، كلما زاد مربع معدل النقود كلما قل احتمال أن يأخذ السند التصنيف عالي الجودة Aa بمقدار 0.02 لكل زيادة بمقدار الوحدة في هذا المعدل.

مثال 3.15:

التنبؤ بعدم دفع السند المالي Predicting bond defaults

للتنبؤ باحتمال عدم دفع تعهدات السند المالي، قام دانيال رابينفيلد بدراسة عينة من 35 بلدية في مقاطعة Massachuzetts في سنة 1930، وقد حدث هناك بالفعل حالات لعدم دفع تعهدات السند المالي نموذج LPM الذي قام باختياره وتقديره هو كالتالي[12]:

$$\hat{P} = \underset{(0.29)}{1.96} - \underset{(0.009)}{0.029}\,TAX - \underset{(2.13)}{4.86}\,INT + \underset{(0.028)}{0.063}\,AV + \underset{(0.003)}{0.007}\,DAV - \underset{(0.88)}{0.48}\,WELF \qquad (2.3.15)$$

$$R^2 = 0.36$$

بحيث إن P=0 إذا قامت البلدية بعدم الدفع و 1 بخلاف ذلك

و TAX= متوسط معدل الضرائب في 1929، 1930 و 1931 .

TNT = نسبة الميزانية الحالية المخصصة لدفع الفوائد في 1930 .

AV = نسبة نمو في تقييم أصول الملكية في خلال الفترة من 1925 إلى 1930 .

DAV = نسبة إجمالي صافي الدين المباشر إلى إجمالي تقدير الأصول في 1930 .

WELF= نسبة الأموال الموجهة في ميزانية 1930 إلى المؤسسات الخيرية والمعاشات ومساعدات الجنود.

تفسير (2.3.15) هو أيضاً تفسير تقليدي. باعتبار ثبات العوامل الأخرى، فإن الزيادة في معدل الضرائب بمقدار دولار واحد سيزيد احتمال عدم سداد السند المالي بحوالي 0.03 أو 3%. قيمة R^2 إذاً صغيرة نوعاً ما كما هو ملاحظ من قبل، في LPM، قيمة R^2 عموماً تميل إلى كونها صغيرة وهي تستخدم بشكل محدود في الحكم على جودة تقدير النموذج.

(12) مرجع D. Rubinfeld, "An econometric Analysis of the Market for General Municipal Bonds," unplished doctoral dissertation, Massachusetts Instiitute of Technology, 1972. The results given in this example art reproduced from Robert S. Pindyck and Daniel L. Rubinfeld, Econometric Models and Economic Forecasts, 2d ed., McGraw-Hill, New York, 1981, p. 279.

4.15 بدائل الـ LPM : ALTERNATIVES OF LPM

كما رأينا الـ LPM تعاني العديد من المشكلات وهي :

(1) عدم اتباع u_i للتوزيع الطبيعي، (2) اختلاف التباين لـ u_i، (3) احتمال أن تقع رقم $\hat{Y}_i$ خارج الحدود 0 و 1، (4) بوجه عام القيم المنخفضة لـ R^2.

ولكن هذه المشكلات من الممكن التغلب عليها. على سبيل المثال، من الممكن استخدام WLS لحل مشكلة اختلاف التباين وزيادة حجم العينة لتقليل مشكلة عدم اتباع الخطأ للتوزيع الطبيعي. وبتعديل قيود طريقة المربعات الصغرى، وأساليب البرامج الرياضية من الممكن عمل الاحتمالات المقدرة التي تقع في حدود الفترة 0 و 1.

لكن بعد ذلك تكون المشكلة الأساسية في الـ LPM هي أنه نموذج غير مقنع، حيث يفرض أن $P_i = E(Y = 1|X)$ تتزايد بشكل خطي مع X، وبالتالي التأثير الحدي لـ X يظل ثابتاً طوال الوقت. وبالتالي مثلاً في مثال امتلاك المنزل السكني وجدنا أن X تتزايد بوحدة واحدة (1000 $) يجعل احتمال امتلاك المنزل السكني يزيد بمقدار ثابت وهو 0.1 وذلك بغض النظر عن كون مستوى الدخل 8.000 $, 10.000 $, 18.000 $ أو 22.000 $ وهذا يعتبر غير واقعي على الإطلاق. في الواقع يتوقع أن تكون P_i مرتبطة بشكل غير خطي مع X_i فمثلاً عندما يكون دخل الأسرة منخفضاً، فإنها لن تمتلك المنزل السكني، ولكن عندما يكون الدخل مرتفعاً نوعاً ما وبشكل كاف، مثلاً X^*، فإنه من المحتمل أن تمتلك الأسرة المنزل السكني، وأي زيادة في الدخل أكبر من X^* سيكون لها تأثير بسيط على احتمال امتلاك الأسرة للمنزل السكني. وبالتالي عند النهايتين الطرفيتين لتوزيع الدخل، احتمال امتلاك المنزل السكني سيكون غير متأثر أفقياً بأي زيادة صغيرة في X.

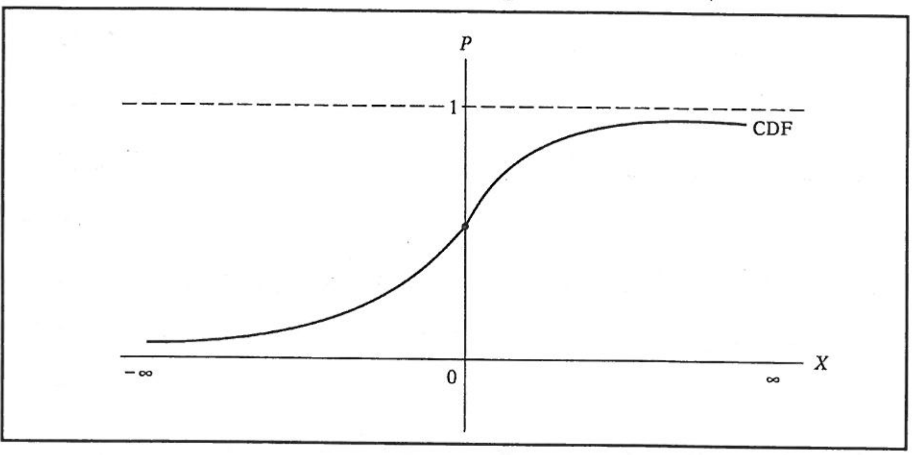
وبالتالي ما نحتاج إليه هو نموذج (احتمالي) يحتوي على هاتين الخاصيتين: (1) كلما زادت X، تزداد $P_i = E(Y = 1|X)$ ولكن بدون أن تتعدى حدود الفترة 0–1 و (2) العلاقة بين P_i و X_i تكون علاقة غير خطية، بمعنى أن تؤول إلى الصفر بمعدل بطيء عندما تكون X صغيرة وتؤول إلى 1 بمعدل بطيء عندما تكون X كبيرة جداً[13].

هندسياً، النموذج المطلوب يأخذ شكلاً قريباً من الشكل (2.15)، لاحظ أن في هذا النموذج الاحتمال يقع بين 0، 1 ويعتبر بشكل غير خطي مع X.

---

(13) مصدر: John Aldrich nd Forrest Nelson, op.cit., p. 26.

القارئ يمكنه أن يدرك أن sigmoid وشكل S في الشكل (2.15) يعبر بشكل ما عن دالة التوزيع التراكمي (CDF) لمتغير عشوائي[14]. وبالتالي من الممكن استخدام CDF كنموذج انحدار عندما يكون المتغير التابع متغيراً ثنائياً. يأخذ القيم 0 و 1. السؤال العملي الآن، كيف يتم اختيار الـ CDF؟ فعلى الرغم من أن جميع الـ CDFs لها الشكل S فلكل متغير عشوائي هناك CDF وحيدة. لأسباب تاريخية وعملية فإن الـ CDF الشائعة الاستخدام للتعبير عن النماذج التي يكون فيها المتغير تابعاً يأخذ القيم 0 و 1.



شكل (2.15) التوزيع التراكمي لمتغير عشوائي

هي (1) التوزيع اللوجيستي و (2) التوزيع الطبيعي. وهذا يعطينا نموذج البروبيت (نورميت). نموذج اللوجيت.

على الرغم من أن الدراسة المفصلة لنماذج اللوجيت والبروبيت خارج نطاق الدراسة في هذا الكتاب، فإننا سنعرض بشكل عام كيف يمكن التقدير باستخدام هذه النماذج، وكيف يمكن تفسير النتائج الخاصة بها.

5.15 نموذج اللوجيت : THE LOGIT MODEL

سنستكمل مثال امتلاك المنزل السكني للحصول على الأفكار الرئيسية لنموذج اللوجيت. نذكر أنه في محاولة تفسير امتلاك المنزل السكني من خلال الدخل كان هناك نموذج LPM كالتالي:

---

(14) كما شرحنا في App.A، الـ CDF الخاصة بالمتغير العشوائي X هي ببساطة احتمال أن يأخذ المتغير العشوائي قيماً أقل من أو تساوي x_0، بحيث إن x_0 هي قيمة عددية محددة لـ X باختصار، $F(X)$، الـ CDF للـ X هي : $F(X=x_0)= P(X\leq x_0)$.

$$P_i = E(Y = 1 \mid X_i) = \beta_1 + \beta_2 X_i \qquad (1.5.15)$$

بحيث إن X هي الدخل، و $Y = 1$ تعني أن الأسرة تمتلك المنزل السكني، والآن دعنا نعتبر المعادلة التالية لتعبير احتمال امتلاك المنزل السكني.

$$P_i = E(Y = 1 \mid X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} \qquad (2.5.15)$$

لسهولة العرض، فإنه يمكن كتابة (2.5.15) كالتالي:

$$P_i = \frac{1}{1 + e^{-Z_i}} = \frac{e^z}{1 + e^z} \qquad (3.5.15)$$

حيث: $Z_i = \beta_1 + \beta_2 X_i$

المعادلة (3.5.15) تمثل ما هو معروف بدالة الاحتمال التراكمية لتوزيع اللوجيستي[15].

من السهل إثبات أن Z_i تأخذ المدى $-\infty$ إلى ∞، P_i تكون بين 0 و 1 وتكون P_i على علاقة غير خطية مع Z_i (بمعنى آخر X_i)، وبالتالي، فإنه مستوفي للشرطين السابق ذكرهما من قبل[16].

ولكن من الواضح أنه باستيفاء هذه الشروط، أوجدنا مشكلة في التقدير، حيث إن P_i ليست غير خطية في X فقط وإنما أيضاً غير خطية في المعاملات β's وهذا واضح من المعادلة (2.5.15) مما يعني أننا لا نستطيع استخدام الـ OLS العادية لتقدير المعالم[17]. ولكن هذه المشكلة تبدو أكثر صعوبة مما هي عليه في واقع الأمر، حيث إن (2.5.15) من الممكن تحويلها إلى الشكل الخطي كالتالي:

إذا كانت P_i هي احتمال أن تمتلك الأسرة المنزل السكني كما هو معطى في (3.5.15) فإن $(1-P_i)$ هو احتمال عدم الامتلاك للمنزل ويمكن التعبير عنه كالتالي:

---

(15) النموذج اللوجيتيكي تم استخدامه بشكل ملحوظ في تحليل النمو مثل نمو السكان، GNP، money supply وهكذا. للقراءة أكثر عن تفاصيل نظرية وعملية عن نماذج اللوجيت والبروبيت، انظر في: مرجع: J.S.Kramer, The Logit Model for Ecomomists, Edward Amold Publishers, London, 1991: and G. S. Maddala, op.cit.

(16) لاحظ أن كلما $Z_i \to +\infty$ فإن e^{-Z_i} تؤول إلى 0 وكلما تؤول $Z_i \to \infty$ فإن e^{-Z_i} تزداد لذكر أن $e = 2.71828$.

(17) بالطبع من الممكن استخدام أساليب التقدير غير الخطية، كما سبق شرحه في الفصل (14)، انظر أيضاً إلى الفقرة 8.15.

$$1 - P_i = \frac{1}{1 + e^{Z_i}} \tag{4.5.15}$$

وبالتالي يمكن أن نقول إن:

$$\frac{P_i}{1 - P_i} = \frac{1 + e^{Z_i}}{1 + e^{-Z_i}} = e^{Z_i} \tag{5.5.15}$$

والآن تكون $\frac{P_i}{1 - P_i}$ هي ببساطة نسبة الأوزان لصالح امتلاك المنزل، أي أنها النسبة بين احتمال أن تمتلك الأسرة المنزل السكني إلى احتمال أنها لن تمتلك هذا المنزل. وبالتالي إذا كان $P_i = 0.8$ فهذا يعني أن الأوزان هي 4 إلى 1 لصالح امتلاك الأسرة للمنزل السكني. والآن إذا أدخلنا اللوغاريتم الطبيعي على المعادلة (5.5.15) فنحصل على نتيجة مهمة جداً وهي:

$$\begin{aligned} L_i = \ln\left(\frac{P_i}{1 - P_i}\right) &= Z_i \\ &= \beta_1 + \beta_2 X_i \end{aligned} \tag{6.5.15}$$

حيث إن L هي Log نسبة الأوزان وهي ليست فقط خطية في الـ X وإنما (من ناحية التقدير) خطية في المعالم أيضاً[18]. L يطلق عليه اللوجيت للنماذج المماثلة لـ (6.5.15).

لاحظ الخصائص التالية لنموذج اللوجيت:

1 - عندما تكون P بين 0 و 1 (بمعنى أن Z تكون بين ∞ -و ∞) فإن اللوجيت L يكون بين ∞- و ∞. بمعنى أنه على الرغم من أن الاحتمالات (بالضرورة) تقع بين 0 و 1، فإن اللوجيت غير محدودة.

2 - بالرغم من أن L خطية في X، الاحتمالات نفسها غير خطية، وهذا الخاصية تتناقض مع النموذج LPM الموجود في (1.5.15) عندما كانت الاحتمالات تتزايد خطياً مع X[19].

3 - على الرغم من أننا أدخلنا قيماً منفردة للمتغير X أو متغير منحدر عليه واحد فقط في النموذج السابق. فإنه من الممكن إضافة العديد من المتغيرات المنحدرة عليها، كما هو موضح في النظرية.

---

(18) تذكر أن فرض التوزيع الطبيعي في الـ OLS لايتطلب أن تكون X بالضرورة خطية، وبالتالي يمكن أن يوجد X_2، X_3 وهكذا كمتغيرات منحدر عليها. في عرضنا السابق العلاقة خطية في المعالم شئ حاسم.

(19) باستخدام الحساب من الممكن إثبات أن $\frac{dp}{dx} = \beta_2 P(L - P)$ الذي يوضح أن معدل التغير في الاحتمال بالنسبة للـ X يتوقف ليس فقط على β_2 ولكن أيضاً على مستوى الاحتمال الذي يقيس التغير (اقرأ أكثر عن ذلك في الفقرة 7.15). ونلاحظ أن تأثير تغير وحدة واحدة في X_i على P يكون أكبر عندما $P= 0.5$ وأقل عندما تكون P قريبة من 0 أو 1.

4- إذا كانت L (اللوجيت) قيمة موجبة، فإن هذا يعني أنه عندما تزداد قيمة المتغير المشارك (أو المتغيرات) المنحدرة عليها، فإنه تزداد أوزان المتغير المنحدر 1 (بمعنى وقوع بعض الأحداث المهتم بدراستها). إذا كانت L قيمة سالبة، فإن أوزان المتغير المنحدر تساوي 1 تقل مع زيادة قيم X. بعبارة أخرى، فإن اللوجيت يصبح سالبًا ويزداد كبر التأثير، وله زيادة كبيرة كلما زادت نسبة الأوزان من 1 إلى مالانهاية[20].

5 – وبصياغة أكثر دقة، فإن تغيير نموذج اللوجيت المعطى في (6.5.15) هي كالتالي: الميل β_2 يقيس التغير في L لكل وحدة واحدة من المتغير في X، بمعنى أنه يعبر عن كيفية تغير لوغاريتم الأوزان لصالح امتلاك منزل الزوجية، كلما تغير الدخل بوحدة واحدة، فمثلاً، 1000 $. الجزء المقطوع من المحور الصادي β_1 هو قيمة لوغاريتم الأوزان لصالح امتلاك المنزل السكني عندما يكون الدخل مساويًا للصفر. مثل العديد من تفسيرات الجزء المقطوع من المحور الصادي، يكون التفسير ليس له معنى واقعي.

6 – بمعلومية مستوى معين من الدخل، مثلا X^*، إذا أردنا تقدير احتمال امتلاك المنزل السكني، وليس تقدير الأوزان لصالح امتلاك المنزل السكني، فهذا يمكن القيام به مباشرة باستخدام (3.5.15) بمجرد توافر التقدير $\beta_1+\beta_2$. وهنا يظهر سؤال في غاية الأهمية وهو: كيف تقدر β_1 و β_2 أولاً؟ الإجابة معطاة في الفقرة التالية.

7 – في حين أن LPM يفترض وجود علاقة خطية بين P_i و X_i، فإن نموذج اللوجيت يفترض وجود علاقة خطية نسبة الأوزان و X_i.

6.15 تقدير نموذج اللوجيت: ESTIMATION OF THE LOGIT MODEL

لغرض التقدير يتم كتابة (6.5.15) كالتالي:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_i + u_i \qquad (1.6.15)$$

سوف نناقش خواص مقدار الخطأ العشوائي u_i قريباً.

لتقدير (1.6.15)، نحتاج (بعض النظر عن قيم X_i) قيم المتغيرات المنحدرة أو اللوجيت L_i. هذا يعتمد بالضرورة على نوع البيانات المتوافرة للتحليل. نحن نفرق بين نوعين من البيانات:

(1) بيانات مفردة (المستوى الجزيئي) Micro level و (2) بيانات تجميعية (مكررة).

---

(20) الفكرة الخاصة بديفد جارسون.

بيانات على مستوى فردي : Data at the individual level

إذا كانت لدينا بيانات عن الأسرة منفردة. كما هو الحال في جدول (1.15) فإن تقديرات OLS للمعادلة (1.6.15) غير ممكنة. وهذا من السهل إثباته. ففي إطار البيانات المعطاة في جدول (1.15)، فإن $1=P_i$ إذا كانت الأسرة تمتلك المنزل السكني و$0=P_i$ إذا كانت لاتمتلك المنزل ولكن إذا عوضنا بهذه القيم مباشرة في اللوجيت L_i نحصل على:

في حالة امتلاك الأسرة للمنزل $$L_i = \ln\left(\frac{1}{0}\right)$$

في حالة عدم امتلاك الأسرة للمنزل $$L_i = \ln\left(\frac{0}{1}\right)$$

نرى بوضوح أن هذه القيم ليس لها معنى، وبالتالي إذا كانت لدينا بيانات على المستوى الفردي أو الجزئي فلا تستطيع تقدير (1.6.15) بطريقة OLS العادية في مثل هذه المواقف قد يكون من الأفضل استخدام **طريقة الإمكان الأعظم** (ML) لتقدير المعالم. وبالرغم من أن مبادئ هذه الطريقة تمت مناقشتها في ملحق الفصل الرابع فإن تطبيقها في الموضوع الحالي ستتم مناقشتها في ملحق 15 A، فقرة 15 1.A من أجل القراء الراغبين في معرفة المزيد عنها[21].

بعض حزم الحاسب الآلي مثل Microsoft، Eviews، Linidep، Shazan، Pcgive و Minitab لديهم أكواد موجودة لتقدير نماذج اللوچيت على مستوى البيانات المنفردة. سنقوم بتوضيح استخدام نموذج ML لاحقاً في هذا الفصل.

بيانات تجميعية أو مكررة : Grouped or Replicated data

الآن دعنا نستعرض البيانات المعطاة في جدول (4.15). هذا الجدول يعطي بيانات على مجموعة من الأسر مجمعة أو مكررة (مشاهدات مكررة) وفقاً لمستوى الدخل وعدد الأسر التي تمتلك المنزل السكني عند هذا المستوى من الدخل. فبالنسبة لكل مستوى من الدخل X_i فإنه يوجد N_i أسرة، n_i الذي يمثل عدد الأسر التي تملك المنزل السكني، بمعنى أن $(n_i \leq N_i)$. وبالتالي إذا حسبنا:

(21) للدراسات عن طريقة الإمكان الأعظم واستخدامها في النموذج اللوجيتي، انظر: John Aldrich and Forest Nelson, op. cit., pp. 49-54. See also, Alfred Demarsi, Logit Modeling: Practical Applications, Sage Publications, Newbury Park, Calif., 1992.

$$\hat{P}_i = \frac{n_i}{N_i} \qquad (2.6.15)$$

فإن هذا هو التكرار النسبي، ومن الممكن استخدامه كتقدير للقيمة الحقيقية P_i المرتبطة بكل X_i. إذا كانت N_i كبيرة بشكل كاف، فإن $\hat{P}_i$ سيكون مقدراً منطقيًا جيدًا لـ P_i[22] وباستخدم القيم المقدرة لـ P_i، يمكن الحصول على تقدير اللوجيت كالتالي:

$$\hat{L}_i = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right) = \hat{\beta}_1 + \hat{\beta}_2 X_i \qquad (3.6.15)$$

جدول (4.15) بيانات افتراضية عن X_i (الدخل)، N_i (عدد الأسر التي لها الدخل X_i) و n_i (عدد الأسر التي تمتلك المنزل السكني)

X (thousands of dollars)	N_i	n_i
6	40	8
8	50	12
10	60	18
13	80	28
15	100	45
20	70	36
25	65	39
30	50	33
35	40	30
40	25	20

هذا التقدير سيكون مقدراً جيداً بشكل كاف للقيمة الحقيقية للوجيت L_i إذا كان عدد المشاهدات N_i لكل X_i كبير بشكل معقول.

باختصار، فإن معلومية بيانات مجمعة أو تكرارية مثل جدول (4.15) من الممكن الحصول على بيانات عن المتغير التابع، اللوجيت، لتقدير النموذج المذكور في المعادلة (1.6.15). هل من الممكن عند إذن تطبيق الـ OLS على (3.6.15) وتقدير المعالم بالطريقة المعتادة؟ الإجابة لا يمكن عمل ذلك بالضبط، حيث إننا حتى الآن لم نتعرض إلى صفات مقدار الخطأ العشوائي. من الممكن إثبات أنه إذا كانت N_i كبيرة نوعاً ما، وكانت كل مفردة موجودة في فئة داخلية X_i موزعة بشكل مستقل كمتغير ذي الحدين، فإن:

$$u_i \sim N\left[0, \frac{1}{N_i P_i(1-P_i)}\right] \qquad (4.6.15)$$

(22) من مبادئ الإحصاء، تذكر أن احتمال الحدث هو نهاية التكرار النسبي عندما يؤول حجم العينة إلى مالا نهاية.

وتكون u_i لهـا التوزيع الطبيـعي بتوقـع صفـر وتبـاين يسـاوي $\frac{1}{N_iP_i(1-P_i)}$ [23]، وبالتالي كـما في حالة LPM، مقـدار الخطأ في نموذج اللوجيت يعـاني من مشكلة اختلاف التباين، وبالتالي بدلاً من استخدام OLS سنحتاج إلى استخدام طريقة المربعات الصغرى المرجحة WLS لأغراض تجريبية يتم استبدال القيمة المجهولة P_i بـ $\hat{P}_i$ وتستخدم:

$$\hat{\sigma}^2 = \frac{1}{N_i\hat{P}_i(1-\hat{P}_i)} \qquad (5.6.15)$$

ويستخدم كتقدير لـ σ^2

دعنا الآن نستعرض الخطوات المختلفة في التقدير للانحدار اللوجيت (1.6.15).

1 - لكل مستوى دخل X، احسب احتمال امتلاك المنزل السكني على أنه: $\hat{P}_i = \frac{n_i}{N_i}$.

2 - لكل X_i، احصل على اللوچيت كالتالي[24]:

$$\hat{L}_i = \ln[\hat{P}_i/(1-\hat{P}_i)]$$

3 - لحل مشكلة اختلاف التباين، حول (1.6.15) كالتالي[25]:

$$\sqrt{w_i}L_i = \beta_1\sqrt{w_i} + \beta_2\sqrt{w_i}X_i + \sqrt{w_i}u_i \qquad (6.6.15)$$

والذي يكتب على الشكل التالي:

$$L_i^* = \beta_1\sqrt{w_i} + \beta_2 X_i^* + v_i \qquad (7.6.15)$$

بحيث إن الأوزان $w_i = N_i\hat{P}_i(1-\hat{P}_i)$، $L_i = \hat{L}_i$ المحولة أو الموزونة، X_i^* قيمة X_i المحولة أو الموزونـة و v_i = مقـدار الخـطأ المحـول. من السهل إثبـات أن مقـدار الخطأ المحول v_i له خـاصيـة ثبات التباين مـع الوضـع في الاعتبار أن التبايـن الأصـلي للخـطأ هو $\sigma_u^2 = 1/[N_iP_i(1-P_i)]$.

---

(23) كما هو موضح في مقدمة نظرية الاحتمالات، $\hat{P}_i$، احتمال النجاح (في المثال يكون امتلاك المنزل السكني) يتبع توزيع ذي الحدين بتوقع يساوي القيمة الحقيقية P_i وتباين يساوي $P_i(1-P_i)/N_i$ وبالتالي كلما زادت N_i كلما أمكن تقريب توزيع ذي الحدين إلى التوزيع الطبيعي. الخصائص التوزيعية لـ u_i المعطى في (4.6.15) تنبع من هذه النظرية الأساسية. للتفاصيل انظر Henry Theil, "On the Relationships Involving Qualitative Variables," American Journal of Sociology, vol. 76, July 1970. pp. 103-154

(24) بما أن $\hat{P}_i \frac{n_i}{N_i}$، من الممكن أن يتم التعبير عن L_i كالتالي $\hat{L}_i = \ln n_i/(N_i - n_i)$ ومن المهم ملاحظة أنه لتجنب أن تأخذ $\hat{P}_i$ القيمة 0 أو 1 في الناحية العملية فإن $\hat{L}_i$ يقاس كالتالي $\hat{L}_i = \ln(n_i+\frac{1}{2})/(N_i-n_i+\frac{1}{2})$. هذا موصى به للتأكد من أن N_i لن تكون أقل من 5 لكل مستوى X. لتفاصيل أكثر، اقرأ D. R. Cox, Analysis of Binary Data, Methuen, London, 1970, p. 33.

(25) إذا قدرنا (1.6.15) متجاهلين اختلاف التباين، المقدرات على الرغم من عدم تحيزها لن تكون كافية كما نعرف من الفصل (11).

4 - قدر (6.6.15) باستخدام OLS- تذكر أن WLS هي OLS على البيانات المحولة، لاحظ أنه في (6.6.15) لايوجد جزء مقطوع من المحور الصادي مقدم صراحة في النموذج (لماذا؟). وبالتالي سيحتاج الفرد إلى استخدام الانحدار المار خلال نقطة الأصل لتقدير (6.6.15).

5 - كون فترات الثقة أو اختبارات الفروض بالطريقة العادية في إطار OLS ولكن ضع في الاعتبار أن كل الاستنتاجات ستكون سليمة فقط، بشرط أن يكون حجم العينة كبيراً بشكل معقول (لماذا؟). وبالتالي في العينات صغيرة الحجم، نتائج التقديرات لابد من تفسيرها بحذر شديد.

7.15 نموذج اللوجيت المُجمع (GLOGIT) :
THE GROUPED LOGIT (GLOGIT) MODEL

مثال رقمي A Numeniceal Example

لشرح النظرية السابق ذكرها، سنستخدم البيانات المعطاة في جدول (4.15)، وبما أن البيانات في الجدول هي مجموعات، فإن نموذج اللوجيت المستخدم لهذه البيانات سيسمى نموذج اللوجيت المُجمع للاختصار glogit جلوجيت. البيانات الأصلية المطلوبة وحسابات أخرى مرتبطة بالبيانات ومطلوبة لاستخدام glogit موجودة في جدول (5.15). نتائج طريقة المربعات الصغرى المرجحة (7.6.15) المعتمدة على البيانات المعطاة في جدول (5.15) هي كالتالي: لاحظ أنه لايوجد جزء مقطوع من المحور الصادي في (7.6.15) وبالتالي الانحدار المار بنقطة الأصل يعتبر ملائماً في هذه الحالة.

$$\begin{aligned} \hat{L}_i^* &= \quad -1.59474\sqrt{w_i} + \quad 0.07862X_i^* \\ \text{se} &= \quad (0.11046) \qquad\qquad (0.00539) \qquad\qquad (1.7.15) \\ t &= (-14.43619) \qquad (14.56675) \qquad R^2 = 0.9642 \end{aligned}$$

الـ R^2 هي مربع معامل الارتباط بين القيم الحقيقية والمقدرة لـ L_i^*. L_i^* و X_i^* هي القيم المرجحة لـ L_i و X_i كما هو موضح في (6.6.15).

تفسير نموذج اللوجيت المقدر: Interpretation of the estimated logit model

كيف نفسر (1.7.15)؟ هناك طرق عديدة لذلك بعضها بسيط والآخر ليس بسيطاً:

تفسير اللوجيت: كما توضح (1.7.15)، القيمة المقدرة للميل، نقترح أنه لكل وحدة (1000 $) زيادة في الدخل المرجح، فإن اللوغاريتم المرجح للأوزان لصالح امتلاك المنزل السكني يزيد بمقدار 0.08 وحدة. هذه ترجمة حرفية، وبوجه عام ليست مقنعة إلى حد كبير.

جدول (5.15) بيانات لتقدير نموذج اللوجيت للامتلاك

X (thousands of dollars) (1)	N_i (2)	n_i (3)	$\hat{P}_i$ (4) = (3) ÷ (2)	$1-\hat{P}_i$ (5)	$\frac{\hat{P}_i}{1-\hat{P}_i}$ (6)	$\hat{L}_i = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right)$ (7)	$N_i\hat{P}_i(1-\hat{P}_i) = w_i$ (8)	$\sqrt{w_i} = \sqrt{N_i\hat{P}_i(1-\hat{P}_i)}$ (9) = $\sqrt{(8)}$	$\hat{L}_i^* = \hat{L}_i\sqrt{w_i}$ (10) = (7)(9)	$\hat{X}_i^* = X_i\sqrt{w_i}$ (11) = (1)(9)
6	40	8	0.20	0.80	0.25	−1.3863	6.40	2.5298	−3.5071	15.1788
8	50	12	0.24	0.76	0.32	−1.1526	9.12	3.0199	−3.4807	24.1592
10	60	18	0.30	0.70	0.43	−0.8472	12.60	3.5496	−3.0072	35.4960
13	80	28	0.35	0.65	0.54	−0.6190	18.20	4.2661	−2.6407	55.4593
15	100	45	0.45	0.55	0.82	−0.2007	24.75	4.9749	−0.9985	74.6235
20	70	36	0.51	0.49	1.04	0.0400	17.49	4.1825	0.1673	83.6506
25	65	39	0.60	0.40	1.50	0.4054	15.60	3.9497	1.6012	98.7425
30	50	33	0.66	0.34	1.94	0.6633	11.20	3.3496	2.2218	100.4880
35	40	30	0.75	0.25	3.0	1.0986	7.50	2.7386	3.0086	95.8405
40	25	20	0.80	0.20	4.0	1.3863	4.00	2.000	2.7726	80.0000

تفسير الأوزان : Odds Interpretation

تذكر أن $L_i = \ln[P_i/(1-P_i)]$ ، وبالتالي بإدخال الدالة العكسية للوغاريتم على تقدير اللوجيت نحصل على $P_i/(1-P_i)$ وهذا هو نسبة الأوزان. وبالتالي إدخال الدالة العكسية للوغاريتم على (1.7.15) نحصل على :

$$\frac{\hat{P}_i}{1-\hat{P}_i} = e^{-1.59474\sqrt{w_i}+0.07862X_i^*} \qquad (2.7.15)$$
$$= e^{-1.59474\sqrt{w_i}} \cdot e^{0.07862X_i^*}$$

باستخدام الآلة الحاسبة نستطيع بسهولة إثبات أن $e^{0.07862} = 1.0817$ وهذا يعني أنه لكل وحدة زيادة في الدخل المرجح تزداد الأوزان المرجحة لصالح امتلاك المنزل السكني بمقدار 1.0817 أو تقريباً 8.17%. بوجه عام، إذا أدخلت الدالة العكسية للوغاريتم على معامل الانحدار (الميل) J (في حالات وجود أكثر من متغير منحدر عليه في النموذج) وطرحت 1 من المقدار، ثم ضربت النتيجة في 100، سوف نحصل على نسبة التغير في الأوزان بالنسبة لوحدة زيادة واحدة في المتغير المنحدر عليه J.

إذا أردت بشكل عرضي إجراء التحليل على اللوجيت غير المرجح، كل ما تحتاج إلى فعله هو قسمة تقدير الـ L_i^* على $\sqrt{w_i}$. جدول (6.15) يعطي تقديرات اللوجيت المرجحة وغير المرجحة لكل مفردة، بالإضافة إلى بعض البيانات التي سيتم مناقشتها لاحقاً.

حساب الاحتمالات : Computing probabilities

بما أن مصطلحات اللوجيت، ونسبة الأوزان، قد تكون غير معتادة بالنسبة للبعض، فإننا دائماً نستطيع حساب احتمال امتلاك المنزل السكني بالنسبة لمستوى معين من الدخل. افترض أننا نريد حساب مثل هذا الاحتمال عند $X = 20(\$29000)$. بالتعويض عن هذه القيمة في (1.7.15) نحصل على $L_i^* = 0.09311$ وبقسمة ذلك على $\sqrt{w_i} = 4.2661$ (انظر جدول 5.15) نحصل على $\hat{L}_i = 0.02226$ وبالتالي عند مستوى دخل 20.000 $ يكون لدينا :

$$-0.02226 = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right)$$

جدول (6.15) L star ، X star ، L star المقدرة ، الاحتمال ، التغير في الاحتمال (*)

Lstar	Xstar	ELstar	Logit	Probability, $\hat{P}$	Change in probability†
−3.50710	15.1788	−2.84096	−1.12299	0.24545	0.01456
−3.48070	24.15920	−2.91648	−0.96575	0.27572	0.01570
−3.48070	35.49600	−2.86988	−0.80850	0.30821	0.01676
−2.64070	55.45930	−2.44293	−0.57263	0.36063	0.01813
−0.99850	74.62350	−2.06652	−0.41538	0.39762	0.01883
0.16730	83.65060	−0.09311	−0.02226	0.49443	0.01965
1.60120	98.74250	1.46472	0.37984	0.59166	0.01899
2.22118	100.48800	2.55896	0.76396	0.68221	0.01704
3.00860	95.84050	3.16794	1.15677	0.76074	0.01431
2.77260	80.00000	3.10038	1.55019	0.82494	0.01135

وبالتالي :

$$\frac{\hat{P}}{1-\hat{P}_i} = e^{-0.02226} = 1.0225$$

وبحل ذلك نحصل على :

$$\hat{P}_i = \frac{e^{-0.02226}}{1+e^{-0.02226}}$$

يستطيع القارئ إثبات أن تقدير الاحتمال هو 0.4944 . وبالتالي ، فإنه بالنسبة لمستوى الدخل 20.000 $ فإن احتمال امتلاك المنزل السكني هو تقريباً 49 في المائة . جدول (6.15) يوضح الاحتمالات المحسوبة بنفس الطريقة وفقاً لمستويات مختلفة من الدخول. وكما يوضح هذا الجدول، فإن احتمال امتلاك المنزل السكني يزداد مع الدخل، ولكن ليس بشكل خطي كما هو الحال في نموذج LPM .

حساب معدل التغير في الاحتمال: Computing the rate of change of probability

من جدول (6.15) تستطيع أن ترى أن احتمال امتلاك المنزل السكني يعتمد على مستوى الدخل. كيف نستطيع حساب معدل التغير في الاحتمالات عندما يتغير الدخل؟ كما سبق الذكر في الملحوظة رقم 19، فإن هذا يعتمد ليس فقط على تقدير معامل الانحدار (الميل) β_2 فقط ، إنما على مستوى الاحتمال الذي يقاس من خلاله التغير وأيضاً يعتمد على مستوى الدخل المحسوب عنده الاحتمال. لتوضيح ذلك، دعنا نفترض أننا نريد قياس التغير، في احتمال امتلاك المنزل السكني عند مستوى

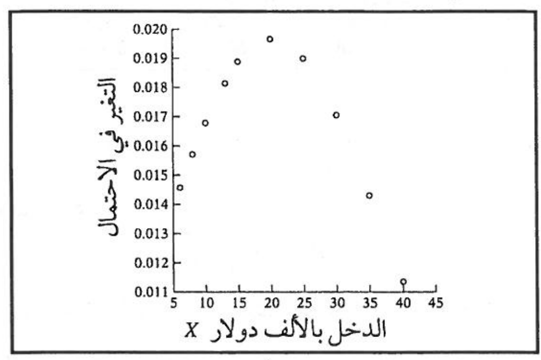
(*) L star و X star من جدول (5.15)، EL star هي القيمة المقدرة L star، لوجيت هو اللوجيت غير المرجح. الاحتمال هو تقدير احتمال امتلاك المنزل السكني. التغير في الاحتمال هو التغير المحسوب لكل وحدة تغير في الدخل

+ محسوبة من $\hat{\beta}_2 \hat{P}(1-\hat{P}) = 0.7862\ \hat{P}(1-\hat{P})$

دخل يساوى 20,000$ ، وبالتالي من الملاحظة 19 ، فإن التغير في الاحتمال لكل وحدة زيادة في الدخل عند مستوى الدخل 20 ألفًا هو :

$$0.01965 = \hat{\beta}\ \ (1-\hat{P}\,)\,\hat{P}\ \ = 0.07862\ (0.5056)\ (0.4944)$$

متروك للقارئ كتمرين ، أن يثبت أنه عند مستوى دخل مساو 40.000$ التغير في الاحتمال هو 0.01135 . جدول (6.15) يوضح التغير في احتمال امتلاك المنزل السكني وفقاً لمستويات مختلفة من الدخول . هذه الاحتمالات أيضاً موضحة في شكل (3.15) .



شكل (3.15) التغير في الاحتمال وعلاقته بالدخل

كتلخيص لمناقشتنا بخصوص نموذج الچي لوجيت glogit دعنا نستعرض نتائج مثال امتلاك المنزل السكني وفقاً لطريقة OLS أو الانحدار غير المرجح

$$\begin{aligned} \hat{L}_i &= \quad -1.6587 \ + \quad 0.7092X_i \\ \text{se} &= \quad (0.0958) \qquad (0.0041) \qquad\qquad\qquad (3.7.15) \\ t &= (-17.32) \qquad (19.11) \qquad\qquad r^2 = 0.9786 \end{aligned}$$

ونترك للقارئ مقارنة هذا الانحدار مع انحدار المربعات الصغرى المرجح والمعطى في (1.7.15) .

8.15 نموذج اللوجيت للبيانات الفردية أو غير المجمعة :
THE LOGIT MODEL FOR UNGROUPED OR INDIVIDUAL DATA

للبدء في الحديث ، دعنا نستعرض البيانات المعطاة في جدول (7.15) دع Y = 1 إذا كانت درجة الطالب النهائية في مادة اقتصاد جزئي متوسطة المستوى هي A و Y = 0 إذا كانت الدرجة النهائية هي B أو C . سبيكتور ومازيو Spector and Mazzeo استخدما نقطة متوسطة الدرجة GAP ، TUCE . والنظام الشخصي للتعليمات PSI تم

استخدامه كمتغير مفسر للدرجة. في هذا المثال نموذج اللوجيت يمكن كتابته كالتالي:

$$L_i = \left(\frac{P_i}{1 - P_i}\right) = \beta_1 + \beta_2 \text{GPA}_i + \beta_3 \text{TUCE}_i + \beta_4 \text{PSI}_i + u_i \qquad (1.8.15)$$

جدول (7.15) بيانات عن تأثير النظام الشخصي للتعليمات (PSI) على درجة المادة العلمية

Observation	GPA grade	TUCE grade	PSI	Grade	Letter grade	Observation	GPA grade	TUCE grade	PSI	Grade	Letter grade
1	2.66	20	0	0	C	17	2.75	25	0	0	C
2	2.89	22	0	0	B	18	2.83	19	0	0	C
3	3.28	24	0	0	B	19	3.12	23	1	0	B
4	2.92	12	0	0	B	20	3.16	25	1	1	A
5	4.00	21	0	1	A	21	2.06	22	1	0	C
6	2.86	17	0	0	B	22	3.62	28	1	1	A
7	2.76	17	0	0	B	23	2.89	14	1	0	C
8	2.87	21	0	0	B	24	3.51	26	1	0	B
9	3.03	25	0	0	C	25	3.54	24	1	1	A
10	3.92	29	0	1	A	26	2.83	27	1	1	A
11	2.63	20	0	0	C	27	3.39	17	1	1	A
12	3.32	23	0	0	B	28	2.67	24	1	0	B
13	3.57	23	0	0	B	29	3.65	21	1	1	A
14	3.26	25	0	1	A	30	4.00	23	1	1	A
15	3.53	26	0	0	B	31	3.10	21	1	0	C
16	2.74	19	0	0	B	32	2.39	19	1	1	A

ملاحظات: الدرجة Y=1 إذا كانت الدرجة النهائية A

0 = إذا كانت الدرجة النهائية B أو C

TUCE= درجة امتحان يعطى في بداية العام الدراسي لمعرفة المعلومات العامة عن الاقتصاد الجزئي

PSI=1 إذا استخدمت طريقة جديدة للتعليم

0 = بخلاف ذلك

GPA نقطة متوسط الدرجة

المصـــدر: L.Spector and M.Mazzeo, "probit analysis and economic education". Journal of: Economic Education, vol. 11, 1980, pp. 37-44.

كما سبق وذكرنا في فقرة 6.15 تستطيع ببساطة جعل P_i =1 إذا كانت الأسرة تمتلك منزلاً سكنياً أو صفراً إذا كانت لا تمتلك المنزل. وفي هذه الحالة، تكون كل من OLS أو المربعات الصغرى المرجحة غير فعالة. لابد من التعامل مع طريقة تقدير غير خطية باستخدام طريقة الإمكان الأعظم. تفاصيل هذه الطريقة معطاة في ملحق A 15، فقرة 15 A.1. وبما أن معظم حزم التحليل الإحصائي الحديثة تشتمل على أكواد خاصة بتقدير نماذج اللوجيت للبيانات غير المجمعة، فإننا سوف نستعرض نتائج النموذج (1.8.15) باستخدام البيانات المعطاة في جدول (7.15) ونتناول كيفية تفسير

مثل هذه النتائج. النتائج معطاة في جدول (8.15) في شكل جدول تم الحصول عليه باستخدام Eviews 4. قبل تفسير النتائج دعنا نستعرض هذه الملاحظات العامة:

1 - بما أننا نستخدم طريقة الإمكان الأعظم، وهي بوجه عام طريقة للعينات كبيرة الحجم، فإن تقديرات الأخطاء القياسية تقاربية (تؤول إلى توزيع ما).

2 - كنتيجة لذلك، فإنه بدلاً من استخدام الإحصاء t لاختبار المعنوية الإحصائية لمعامل ما، فإننا نستخدم الإحصاء (الطبيعي القياسي) Z. وبالتالي الاستدلال قائم على جدول التوزيع الطبيعي، وتذكر أنه إذا كان حجم العينة كبيراً إلى حد ما، فإن توزيع t يؤول إلى التوزيع الطبيعي.

3 - كما سبق، أن ذكرنا، فإن الاختبار التقليدي لجودة التوفيق R^2 لايصلح بشكل عملي في حالة النماذج ذات المتغير المنحدر الثنائي. مقاييس أخرى شبيهة بـ R^2 تسمى pseudo R^2 من الممكن استخدامها إلى جانب العديد من مثل هذه الاختبارات[26]. Eviews تستخدم فيها أحد هذه المقاييس وهو Mac Fadden R^2 ويرمز له بالرمز R^2_{McF}.

جدول (8.15) نتائج انحدار (1.8.15)

Variable	Coefficient	Std. error	Z statistic	Probability
C	-13.0213	4.931	-2.6405	0.0082
GPA	2.8261	1.2629	2.2377	0.0252
TUCE	0.0951	0.1415	0.67223	0.5014
PSI	2.3786	1.0645	2.2345	0.0255
	McFadden $R^2 = 0.3740$	LR statistic (3 df) = 15.40419		

المتغير التابع: الدرجة
الطريقة: اللوجيت الثنائي باستخدام ML
النتائج تم التوصل إليها بعد 5 تكرارات

(26) لدراسات أخرى في هذا الموضوع انظر:
J. Scott long, Regression models for categorical and limited department vauicebles sage publications, New bury park, California, 1997, pp. 102.113.

(27) بشكل فني يتم تعريف ذلك على أنه: $1-(LLF_{ur}/LLF_r)$ بحيث إن LLFur هو اللوغاريتم غير المقيد لدالة الإمكان عندما تكون كل المتغيرات المنحدرة موجودة في النموذج و LLF_r هو اللوغاريتم المقيد لدالة الإمكان عندما يكون الجزء المقطوع من المحور الصادي هو الوحيد الموجود في النموذج. بمعنى أن LLF_{ur} مماثل لـ RSS و LLF_r مماثل لـ TSS في نموذج الانحدار الخطي.

والذي يأخذ القيمة 0.3740 في المثال الحالي(27). R^2 تقع أيضاً بين 0 و 1 مثل R^2. مقياس آخر بسيط لجودة التوفيق هو R^2_{McF} العددي (count R^2) ويعرف على أنه:

$$\text{Count } R^2 = \frac{\text{number of correct predictions}}{\text{total of observations}} \tag{2.8.15}$$

وبما أن المتغير المنحدر عليه في نموذج اللوجيت يأخذ القيمة 1 أو 0، فإذا كان الاحتمال المتنبأ به أكبر من 0.5 نعطي ذلك القيمة 1، ولكن إذا كان أقل من 0.5 فنعطي ذلك القيمة 0. وبعد ذلك نعد المرات التي حدث فيها تنبؤ صحيح، ونحسب R^2 كما هو معطى في (2.7.15)، سنقوم بشرح ذلك قريباً. يجب ملاحظة أنه في نماذج الانحدار الثنائية، فإن جودة التوفيق لها أهمية ثانوية. فالمهم هو الإشارات المتوقعة لمعاملات الانحدار والمنضوية الإحصائية أو العملية لهم.

4 - لاختبار الفرض العدمي الخاص بمعاملات الانحدار (الميل) وتساويها آنيًا بالصفر، الاختبار المماثل لاختبار F في نموذج الانحدار الخطي هو إحصاء نسبة الإمكان (LR). بافتراض صحة الغرض العدمي، فإن الإحصاء LR يتبع توزيع X^2 بدرجات حرية مساوية لعدد المتغيرات المفسرة، ثلاثة متغيرات في مثالنا الحالي (لاحظ أن الجزء المقطوع من المحور الصادي لايدخل في حساب درجات الحرية).

دعنا الآن نفسر نتائج الانحدار في (1.8.15). كل معامل انحدار (الميل) في المعادلة هو معامل انحدار جزيئي ويقيس التغير في اللوجيت المقدر لكل تغير بمقدار الوحدة في قيمة المتغيرات المنحدرة المعطاة (بافتراض ثبات باقي المتغيرات المنحدرة). وبالتالي معامل GPA المساوي لـ 2.8261 يعني بافتراض ثبات باقي العوامل، فإنه إذا زاد الـ GPA بوحدة واحدة ،فإنه في المتوسط يزداد تقدير اللوجيت بمقدار 2.83 وحدة، وبالتالي هناك علاقة موجبة (طردية) بين المتغيرات. وكما ترى، فإن كل المتغيرات المنحدرة الأخرى لها تأثير موجب على اللوجيت، على الرغم من أن تأثير TUCE ليس معنوياً، عموماً فإن كل المتغيرات المنحدرة جميعاً لها تأثير معنوي على الدرجة النهائية، بدليل أن إحصاء LR يساوي 15.4 والمرتبطة بقيمة P(P-value) حوالي 0.0015 والتي تعتبر قيمة صغيرة جداً.

كما سبق وذكرنا ، فإن التفسير الأفضل يكون من خلال الأوزان، والتي يتم الحصول عليها عن طريق استخدام دالة اللوغاريتم العكسية لمعاملات الانحدار المختلفة، وبالتالي إذا أخذنا دالة اللوغاريتم العكسية لمعامل انحدار PSI والذي يساوي 2.3786 سنحصل على ($10.7897 (\approx e^{2.3786})$) وهذا يعني أن الطلاب الذين يدرسون بطرق تعليم حديثة أكثر من 10 مرات في إمكانية الحصول على تقدير A من الطلاب الذين لم يدرسوا طرق التعليم الحديثة مع افتراض ثبات باقي العوامل الأخرى.

افترض أننا نريد حساب الاحتمال الفعلي لأن يحصل الطالب على تقدير A، دعنا نأخذ الطالب رقم 10 في جدول (7.15) ونضع بيانات هذا الطالب الفعلية في نموذج اللوجيت المقدر والمعطى في جدول (8.15). يمكن للقارئ أن يتأكد من أن قيمة اللوجيت المقدرة لهذا الطالب هي 0.8178.

باستخدام المعادلة (2.5.15) يمكن للقارئ بسهولة أن يتأكد من أن القيمة المقدرة للاحتمال هي .0.69351 بما أن الدرجة النهائية الفعلية لهذا الطالب هي A، ونموذج اللوجيت في هذا المثال يعطي احتمال 1 للطالب الذي يحصل على A، فإن القيمة المقدرة للاحتمال وهي 0.69351 ليس بالضبط 1 ولكن قريبة لها.

تذكر الـ count R^2 المعرفة من قبل. جدول 9.15 يعطي القيمة الفعلية والمقدرة للمتغير المنحدر عليه في مثالنا التوضيحي. من هذا الجدول، يمكن أن نلاحظ أنه من 32 مشاهدة توجد 6 مشاهدات خاطئة التنبؤ (الطلاب رقم 14، 19، 24، 26، 31 و32) وبالتالي الـ count R^2 تساوي $0.8125 = 26/ 32$ في حين أن Mc Fadden R^2 تساوي .0.3740 وعلى الرغم من أنه لا يمكن مقارنة هاتين القيمتين بشكل مباشر، إلا أنهما يعطيان فكرة عن مدى التأثير الموجود. بالإضافة إلى أنه ليس من المفروض إعطاء قياس جودة التوفيق أهمية أكثر من اللازم في إطار النماذج التي يكون فيها المنحدر عليه متغيراً مزدوجًا.

جدول (9.15) القيم المقدرة والفعلية الخاصة بالانحدار الموجود في جدول (8.15)

Observation	Actual	Fitted	Residual	Residual plot
1	0	0.02658	−0.02658	
2	0	0.05950	−0.05950	
3	0	0.18726	−0.18726	
4	0	0.02590	−0.02590	
5	1	0.56989	0.43011	
6	0	0.03486	−0.03486	
7	0	0.02650	−0.02650	
8	0	0.05156	−0.05156	
9	0	0.11113	−0.11113	
10	1	0.69351	0.30649	
11	0	0.02447	−0.02447	
12	0	0.19000	−0.19000	
13	0	0.32224	−0.32224	
*14	1	0.19321	0.80679	
15	0	0.36099	−0.36099	
16	0	0.03018	−0.03018	
17	0	0.05363	−0.05363	
18	0	0.03859	−0.03859	
*19	0	0.58987	−0.58987	
20	1	0.66079	0.33921	
21	0	0.06138	−0.06138	
22	1	0.90485	0.09515	
23	0	0.24177	−0.24177	
*24	0	0.85209	−0.85209	
25	1	0.83829	0.16171	
*26	1	0.48113	0.51887	
27	1	0.63542	0.36458	
28	0	0.30722	−0.30722	
29	1	0.84170	0.15830	
30	1	0.94534	0.05466	
*31	0	0.52912	−0.52912	
*32	1	0.11103	0.88897	

(*) تنبؤ غير صحيح

9.15 نموذج البروبيت : THE PROBIT MODEL

كما سبق وذكرنا لتفسير تباين المتغير التابع المزدوج، لابد من استخدام دالة توزيع تراكمي CDF مناسبة. في نموذج اللوجيت يتم استخدام دالة اللوجيستيك التراكمية كما هو موضح في (2.5.15). ولكن هذه ليست هي دالة التوزيع التراكمي الوحيدة المستخدمة. في بعض التطبيقات تستخدم دالة التوزيع التراكمي للتوزيع الطبيعي. النموذج المقدر الذي يعتمد على دالة التوزيع الطبيعي التراكمية(28) معروف باسم نموذج البروبيت، وأحياناً يعرف باسم نموذج النورميت. كفكرة رئيسية

(28) انظر ملحق A لدراسة دالة التوزيع الطبيعي التراكمية. باختصار إذا كان المتغير X يتبع التوزيع الطبيعي بتوقع u وتباين σ^2 فإن دالة التوزيع الاحتمالي هي :

$$f(X) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(X-\mu)^2/2\sigma^2}$$

وتكون دالة التوزيع التراكمي هي :

$$F(X) = \int_{-\infty}^{X_0} \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(X-\mu)^2/2\sigma^2}$$

بحيث إن X_0 هي قيمة محددة لـ X.

من الممكن استبدال CDF اللوجيستيك بـ CDF الطبيعي في معادلة (2.5.15) و اتباع المنهج المذكور في فقرة 5.16. بدلاً من استخدام هذه الطريقة سنقوم بتقديم نموذج البروبيت معتمدين على نظرية المنفعة أو الاختيار النظري الرشيد في السلوك والمقدم من Mc fadden[29].

لاستعراض نموذج البروبيت، افترض أنه في مثال امتلاك المنزل السكني، فإن قرار امتلاك الأسرة I المنزل السكني من عدمه، يعتمد على مؤشر منفعة غير مشاهد I_i (ويعرف أيضاً باسم المتغير المستتر) والذي يحدد من خلال متغير أو أكثر من المتغيرات المفسرة، مثلاً الدخل X_i بحيث إنه كلما زادت قيمة المؤشر I_i زاد احتمال امتلاك الأسرة للمنزل السكني. نعبر عن هذا المؤشر I_i كالتالي:

$$I_i = \beta_1 + \beta_2 X_i \qquad (1.9.15)$$

بحيث إن X_i هو دخل الأسرة رقم i.

كيف يرتبط هذا المؤشر (غير المشاهد) بالقرار الفعلي الخاص بامتلاك المنزل السكني؟ كما سبق دع Y=1 إذا كانت الأسرة تمتلك المنزل و Y=0 إذا كانت لا تمتلكه. والآن من المعقول افتراض أن هناك مستوى حرجًا أو مبدئياً لهذا المؤشر يسمى I_i^* بحيث إنه إذا زاد I_i عن I_i^* فإن الأسرة ستمتلك المنزل السكني، وبخلاف ذلك لن تمتلكه. المستوى المبدئي I_i^* مثل I_i غير مشاهد، ولكن إذا افترضنا أنه يتبع التوزيع الطبيعي بنفس التوقع والتباين فمن الممكن ليس فقط تقدير معالم هذا المؤشر المعطاة في (1.9.15) ولكن أيضاً من الممكن الحصول على بعض المعلومات عن المؤشر غير المشاهد نفسه، وهذا الحساب يتم كالتالي:

بافتراض التوزيع الطبيعي، فإن احتمال أن يكون I_i^* أقل من أو يساوي I_i من الممكن التعبير عنه في صورة CDF التوزيع الطبيعي القياسي كالتالي[30]:

$$P_i = P(Y = 1 \mid X) = P(I_i^* \leq I_i) = P(Z_i \leq \beta_1 + \beta_2 X_i) = F(\beta_1 + \beta_2 X_i) \qquad (2.9.15)$$

بحيث إن $P(Y = 1|X)$ تعني احتمال وقوع الحدث بشرط قيم معينة للـ X أو قيم معينة للمتغيرات المفسرة، و Z_i هو متغير له التوزيع الطبيعي القياسي، بمعنى $Z \sim N(0, \sigma^2)$. F هي الـ CDF للتوزيع الطبيعي القياسي والتي تكتب بشكل تفصيلي كالتالي:

---

(29) D.Mc Fadden, "Conditional logit analysis of qualitative choice behavior", in P.Zarembke (ed.), frontiers in Econometics, academic press, New York, 1973.

(30) التوزيع الطبيعي الذي يكون توقعه= 0 وتباينه= 1 معروف باسم التوزيع الطبيعي القياسي (انظر ملحق A).

$$F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I_i} e^{-z^2/2}\, dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1+\beta_2 X_i} e^{-z^2/2}\, dz \qquad (3.9.15)$$

بما أن P تعبر عن احتمال وقوع الحدث، فيكون هنا احتمال امتلاك الأسرة للمنزل السكني مقاساً بالمساحة تحت منحنى التوزيع الطبيعي القياسي من ∞ـ إلى I_i كما هو موضح في شكل (A 4.15).

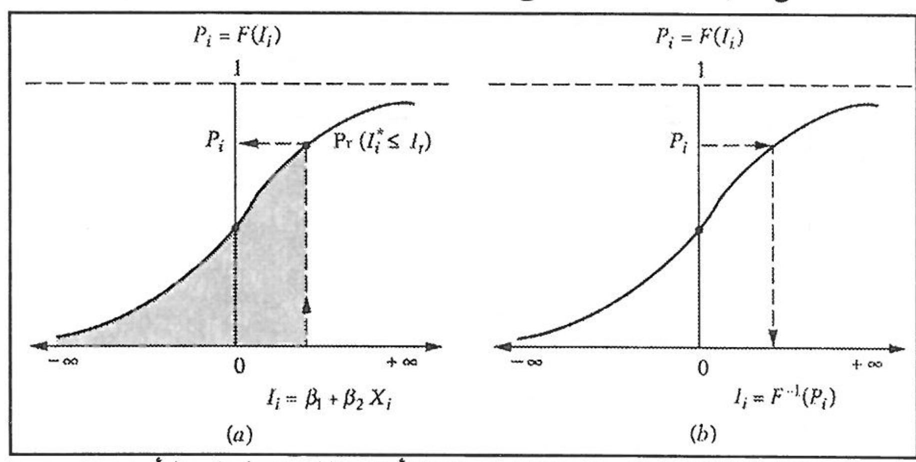
والآن للحصول على معلومات عن I_i، مؤشر المنفعة، والحصول أيضاً على معلومات عن β_1 و β_2 دعنا نأخذ الدالة العكسية لـ (2.9.15) فنحصل على:

$$I_i = F^{-1}(I_i) = F^{-1}(P_i)$$
$$= \beta_1 + \beta_2 X_i \qquad (4.9.15)$$

بحيث إن F-1 هي الدالة العكسية لـ CDF التوزيع الطبيعي. ما يعنيه ذلك من السهل فهمه من الشكل (4.15). في الجزء a من هذا الشكل نحصل على الاحتمال الترتيبي (التراكمي) لامتلاك المنزل السكني، بحيث إن $L_i^* \leq I_i$ في حين أن الجزء b من الشكل نحصل على الإحداثي الأفقي (السيني) للقيمة I_i بمعلومية القيمة P_i وذلك ببساطة هو عكس الذي يحدث عادة.

ولكن كيف بشكل فعلي نحصل على المؤشر I_i والتقديرات β_1 و β_2؟ في حالة نموذج اللوجيت الإجابة تعتمد على ما إذا كانت البيانات مجمعة أو غير مجمعة.

سندرس هاتين الحالتين كلاً على حدة.



شكل (4.15) نموذج البروبيت (a) بمعلومية I_i اقرأ P_i من الإحداثي الصادي الرأسي (b) بمعلومية P_i، اقرأ I_i من الإحداثي السيني (الأفقي)

تقدير البروبيت للبيانات المجمعة: الچي بروبيت:

Probit Estimation with Grouped data: Gprobit

سنستخدم نفس البيانات التي استخدمناها من قبل مع الچي لوجيت g logit والمعطاة في جدول (15.4). بما أنه لدينا فعليًا ، التكرار النسبي (القياس التجريبي للاحتمال) لامتلاك المنزل السكني وفقاً لمستويات مختلفة من الدخل المعطى في جدول (5.15) فمن الممكن استخدامه للحصول على I_i من CDF التوزيع الطبيعي، كما هو موضح في جدول (10.15) أو في شكل (5.15).

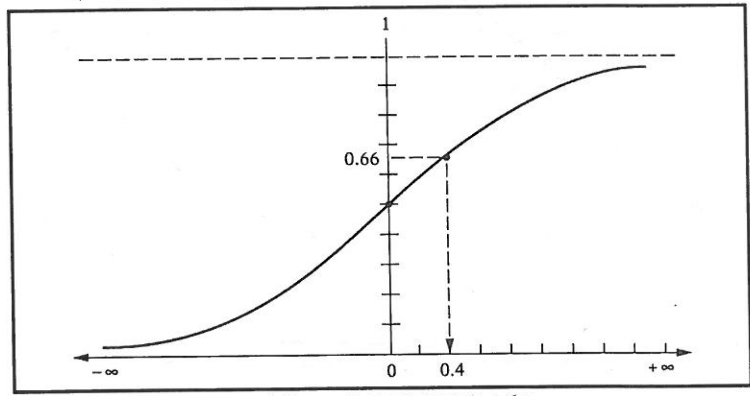
بمجرد حصولنا على قيم I_i المقدرة نحصل على تقديرات β_1 و β_2 بشكل مباشر، كما سنوضح قريباً. وعموماً ضع في الاعتبار أنه في إطار تحليل البروبيت فإن مؤشر المنفعة غير المشاهد I_i معروف باسم الانحراف الطبيعي المتكافئ (n.e.d) أو ببساطة نورميت. بما أن n.e.d أو I_i سيكون سالباً عندما تكون $P_i < 0.5$ في الواقع يتم إضافة العدد 5 إلى الـ n.e.d والنتيجة تسمى بروبيت.

جدول (10.15) تقدير المؤشر I_i من CDF التوزيع الطبيعي القياسي

$\hat{P}_i$	$I_i = F^{-1}(\hat{P}_i)$
0.20	−0.8416
0.24	−0.7063
0.30	−0.5244
0.35	−0.3853
0.45	−0.1257
0.51	0.0251
0.60	0.2533
0.66	0.4125
0.75	0.6745
0.80	0.8416

لاحظ أن : (1) P_i من جدول (5.15)

(2) I_i مقدرة من الـ CDF التوزيع الطبيعي القياسي



شكل (5.15) CDF التوزيع الطبيعي

شرح الجي بروبيت باستخدام المثال المنزلي:

Illustration of Gprobit using housing example

دعنا نستكمل مثالنا المنزلي. لقد قدمنا من قبل نتائج نموذج الجي لوجيت لهذا المثال. أما نتائج البروبيت التجميعي (الجي بروبيت) الخاصة بنفس البيانات فهي كالتالي:

باستخدام الـ n.e.d (=I) المعطى في جدول (10.15)، نتائج الانحدار معطاة في جدول (11.15)[31]. نتائج الانحدار باستخدام البروبيت (5 +n.e.d =) معطاة في جدول (12.15).

استثناء الجزء المقطوع من المحور الصادي، هذه النتائج مطابقة لما هو موجود في الجدول السابق. ولكن ليس من المفروض أن يكون ذلك مستغرباً (لماذا؟).

جدول (11.15) المتغير التابع: I

Variable	Coefficient	Std. error	t statistic	Probability
C	-1.0166	0.0572	-17.7473	1.0397E-07
Income	0.04846	0.00247	19.5585	4.8547E-08

$R^2 = 0.97951$ Durbin-Watson statistic = 0.91384

جدول (12.15) المتغير التابع: البروبيت

Variable	Coefficient	Std. error	t statistic	Probability
C	3.9833	0.05728	69.5336	2.03737E-12
Income	0.04846	0.00247	19.5585	4.8547E-08

$R^2 = 0.9795$ Durbin-Watson statistic = 0.9138

لاحظ أن هذه النتائج ليست مصححة من اختلاف التباين (انظر تمرين 12.15)

تفسير تقديرات البروبيت في جدول (11.15):

Interpretation of the probit estimates in table 15.11

كيف نفسر النتائج السابقة؟ افترض أننا نريد معرفة تأثير تغير X بوحدة واحدة (الدخل مقاس بالألف دولار) على احتمال أن Y=1 وهو شراء الأسرة للمنزل السكني. للقيام بذلك، انظر معادلة (2.9.15) نحن نريد الحصول على تفاضل هذه الدالة بالنسبة لـ X (وهذا هو معدل تغير في الاحتمال بالنسبة للدخل). سيكون هذا التفاضل كالتالي:

$$\frac{dP_i}{dX_i} = f(\beta_1 + \beta_2 X_i)\beta_2 \qquad (5.9.15)^{(32)}$$

(31) النتائج المعطاة غير مصححة من اختلاف التباين. انظر تمرين 12.15 للأسلوب الملائم لتصحيح اختلاف التباين.

(32) نستخدم قاعدة السلسلة للتفاضل:

$$\frac{dP_i}{dX_i} = \frac{dF(t)}{dt} \cdot \frac{dt}{dX}$$

بحيث إن: $t = \beta_1 + \beta_2 X_i$

بحيث إن $f(\beta_1 + \beta_2X)$ هي دالة احتمال التوزيع الطبيعي القياسي محسوبة عند القيمة $\beta_1 + \beta_2X$. ويمكنك أن تلاحظ أن هذا الحساب يعتمد على قيمة محددة للمتغير X دعنا نأخذ قيمة للـ X من جدول (5.15) مثلاً $X = 6$ (ألف دولار). باستخدام القيم المقدرة للمعاملات المعطاة في جدول (11.15) نريد الحصول على دالة التوزيع الاحتمالي للتوزيع الطبيعي عند $f[-1.0166 + 0.04846(6)] = f(-0.72548)$ وبالعودة على جداول التوزيع الطبيعي سنجد أنه بالنسبة لـ $Z = -0.72548$ فإن الاحتمال الطبيعي هو 0.3066[33] . والآن بضرب هذه القيمة في تقدير معامل الانحدار المساوي 0.04846 نحصل على 0.01485 . هذا يعني أننا إذا كان لدينا مستوى دخل يساوي 6000 $ وزاد هذا الدخل بمقدار 1000 $ فإن احتمال أن تشتري الأسرة المنزل السكني يرتفع بمقدار 1.4 في المائة. قارن هذه النتائج مع نظيرها الموجود في جدول (6.15).

كما ترى من المناقشة السابقة مقارنة بـ LPM ونماذج اللوجيت ، فإن حساب التغير في الاحتمال باستخدام نموذج البروبيت يعتبر مثلاً نوعاً ما.

بدلاً من حساب التغير في الاحتمال، افترض أنك تريد إيجاد الاحتمالات المقدرة من نموذج الجي بروبيت المقدر. يمكن القيام بذلك بسهولة. باستخدام البيانات الموجودة في جدول (11.15) والتعويض بقيم X من جدول (5.15) يمكن للقارئ أن يتأكد من أن القيم المقدرة لـ n.e.d هي كالتالي:

X	6	8	10	13	15	20	25	30	35	40
Estimated n.e.d.	−0.72	−0.63	−0.53	−0.39	−0.29	−0.05	0.19	0.43	0.68	0.92

الآن الحزم الإحصائية مثل minitab ممكن بسهولة أن تحسب الاحتمالات (التراكمية) المرتبطة بالقيم المختلفة لـ n.e.d's. على سبيل المثال، فإنه بالنسبة لقيمة n.e.d المساوية 0.63 − فإن الاحتمال المقدر هو 0.2647 أما بالنسبة لقيمة n.e.d المساوية لـ 0.43 فإن الاحتمال المقدر هو 0.6691. إذا قارنت هذه التقديرات مع القيم الفعلية المعطاة في جدول (5.15) سوف تجد أنهما قريبان إلى حد كبير، وذلك يجعلنا نعتبر هذا النموذج نموذجاً جيداً. بالرسم ما قمنا بعمله الآن معطي في شكل (4.15).

نموذج البروبيت للبيانات غير التجميعية أو المفردة:

The probit model for ungrouped or individual data

دعنا نعود إلى جدول (7.15) الذي يعطي بيانات عن 32 مفردة خاصة بدرجات الطلبة النهائية في امتحان مادة الاقتصاد الجزئي وعلاقة ذلك بالمتغيرات GPA ، TUCE و PSI. نتائج انحدار اللوجيت معطاة في جدول (8.15) دعنا نرى كيف ستكون

(33) لاحظ أن التوزيع الطبيعي القياسي Z يأخذ المدى ∞− إلى ∞ ولكن دالة الاحتمال $f(Z)$ هي دائماً موجبة.

نتائج نموذج البروبيت. لاحظ أنه في حالة نموذج اللوجيت للبيانات المفردة سنحتاج لاستخدام طريقة تقدير غير خطية معتمدة على طريقة الإمكان الأعظم. نتائج الانحدار المحسوبة باستخدم Eviews 4 معطاة في جدول (13.15).

جدول (13.15)

المتغير التابع: الدرجة

الطريقة: بروبيت ثنائي – ML

النتائج تم الوصول إليها بعد 5 تكرارات

Variable	Coefficient	Std. error	Z statistic	Probability
C	-7.4523	2.5424	-2.9311	0.0033
GPA	1.6258	0.6938	2.3430	0.0191
TUCE	0.0517	0.0838	0.6166	0.5374
PSI	1.4263	5950	2.3970	0.0165

LR statistic (3 df) = 15.5458 McFadden R^2 = 0.3774
Probability (LR stat) = 0.0014

جدول (14.15)

المتغير التابع: الدرجة

Variable	Coefficient	Std. error	t statistic	Probability
C	-1.4980	0.5238	-2.8594	0.0079
GPA	0.4638	0.1619	2.8640	0.0078
TUCE	0.0104	0.0194	0.5386	0.5943
PSI	0.3785	0.1391	2.7200	0.0110

R^2 = 0.4159 Durbin-Watson d = 2.3464 F statistic = 6.6456

نتائج نموذج البروبيت يمكن مقارنتها مع نتائج نموذج اللوجيت كيفياً، بحيث إن الـ GPA و PSI كلاً على حدة يعتبر متغيراً معنوياً إحصائياً في كل من النموذجين وبما أن قيمة إحصاء LR هي 15.5458 وقيمة الـ P-value المصاحبة له هي 0.0014 فإن كل المتغيرات المفسرة معاً تعتبر متغيرات معنوية إحصائياً. ومع ذلك، فإنه لا يمكن مقارنة معاملات الانحدار لنماذج اللوجيت مع نظيرتها لنماذج البروبيت.

للمقارنة تم عرض النتائج الخاصة بنموذج الاحتمال الخطي LPM لبيانات الدرجات النهائية الموجودة في جدول (14.15). مرة أخرى كيفياً تعتبر نتائج LPM مشابهة لنتائج نماذج اللوجيت والبروبيت بحيث إن GPA و PSI كلاً على حدة يعتبر متغيراً معنوياً إحصائياً ولكن TUCE ليس متغيراً معنوياً. وأيضاً كل درجات المتغيرات المفسرة معاً لها تأثير معنوي على الدرجة النهائية بحيث إن قيمة F المحسوبة هي 6.6456 وهي قيمة معنوية إحصائياً نظراً لأن قيمة P-value المصاحبة لها هي 0.0015.

التأثير الحدي على وحدة التغير في قيمة المتغير المنحدر في عدد من نماذج الانحدار المختلفة:
The Marginal Effect on a unit change in the value of a regressor in the various regression models

في نموذج الانحدار الخطي معامل الانحدار (الميل) يقيس التغير في القيمة المتوسطة للمتغير المنحدر عليه لكل تغير بقيمة الوحدة في قيم المتغير المنحدر بافتراض ثبات باقي العوامل الأخرى.

في الـ LPM، معامل الانحدار يقيس مباشرة التغير في احتمال وقوع الحدث كنتيجة للتغيير بمقدار الوحدة في المتغير المنحدر بافتراض ثبات تأثير باقي المتغيرات الأخرى.

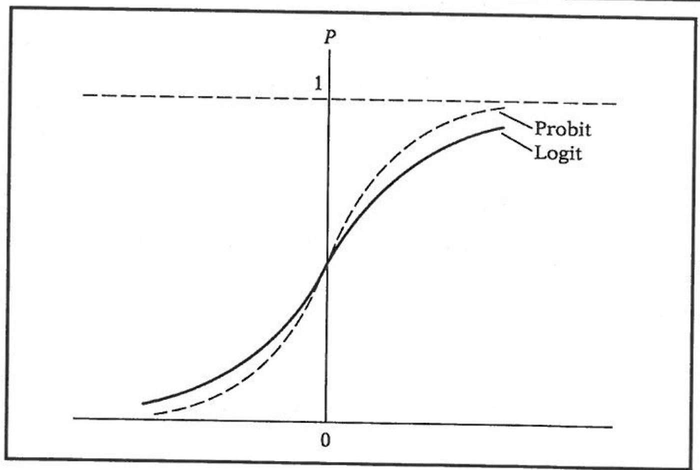
في نموذج اللوجيت معامل الانحدار لمتغير ما يعطي التغير في لوغاريتم الأوزان المرتبطة بالتغير في مقدار الوحدة لهذا المتغير، وأيضاً بافتراض ثبات باقي العوامل (المتغيرات) الأخرى. ولكن كما هو ملاحظ من قبل، فإنه في نموذج اللوجيت معدل التغيير في احتمال وقوع الحدث يساوي $\beta_j P_i (1-P_i)$ بحيث إن β_j هو معامل (الانحدار الجزيئي) للمتغير المنحدر J ولكن لحساب P_i لكل المتغيرات الموجودة في التحليل يتم استخدامها.

في نموذج البروبيت كما سبق ورأينا، معدل التغير في الاحتمال معقد نوعاً ما، ويساوي $\beta_j f(Z_i)$ بحيث إن $f(Z_i)$ هي دالة احتمال التوزيع الطبيعي القياسي و $Z_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki}$ ويكون ذلك هو نموذج الانحدار المستخدم في التحليل.

وبالتالي في كل من نماذج اللوجيت والبروبيت كل المتغيرات المنحدرة تدخل في حساب التغير في الاحتمال، في حين أنه في الـ LPM يدخل فقط المتغير j المرتبط بمعامل الانحدار. وقد يكون هذا هو سبب لكثرة استخدام نموذج LPM في السابق.

10.15 نماذج اللوجيت والبروبيت : LOGIT AND PROBIT MODELS

على الرغم من أنه في مثال الدرجة النهائية LPM واللوجيت والبروبيت أعطت نتائج متماثلة كيفياً سنهتم فقط بنتائج نماذج اللوجيت والبروبيت نظراً للمشكلة الموجودة في الـ LPM والتي سبق ذكرها من قبل بين اللوجيت والبروبيت، من الأفضل؟ في العديد من التطبيقات تكون النماذج متماثلة نوعاً ما. فالفرق الرئيسي أن التوزيع اللوجيتيك متفلطح أكثر قليلاً، وذلك يمكن ملاحظته في الشكل (6.15). بمعنى أن الاحتمال الشرطي لـ P_i يؤول إلى الصفر أو الواحد بمعدل أقل في نموذج اللوجيت عن نموذج البروبيت. وذلك يمكن ملاحظته بوضوح من جدول (15.15) وبالتالي لا يوجد سبب فعلي لتفضيل نموذج على الآخر. في الجانب العملي (التطبيقي) معظم الباحثين يختارون نموذج اللوجيت، حيث إنه أسهل رياضياً.



شكل (6.15) التوزيع التراكمي للوجيت والبروبيت

جدول (15.15) قيم دوال الاحتمال التراكمية

Z	Cumulative normal $P_1(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-s^2/2} ds$	Cumulative logistic $P_2(Z) = \frac{1}{1+e^{-z}}$
−3.0	0.0013	0.0474
−2.0	0.0228	0.1192
−1.5	0.0668	0.1824
−1.0	0.1587	0.2689
−0.5	0.3085	0.3775
0	0.5000	0.5000
0.5	0.6915	0.6225
1.0	0.8413	0.7311
1.5	0.9332	0.8176
2.0	0.9772	0.8808
3.0	0.9987	0.9526

على الرغم من تماثل النموذجين، لابد من أن يتم تفسير معاملات الانحدار بشكل دقيق في كل من النموذجين. على سبيل المثال، في مثال الدرجة النهائية، معامل انحدار GPA المساوي 1.6258 لنموذج البروبيت يساوي 2.8261 في نموذج اللوجيت ولا يمكن مقارنتهما مباشرة. السبب في ذلك أنه على الرغم من أن التوزيع القياسي اللوجيستيكي (أساس اللوجيت) والتوزيع الطبيعي القياسي (أساس البروبيت) كل منهما له توقع يساوي الصفر، فإن تباينهما مختلف: 1 للطبيعي القياسي (كما هو معروف) و $\pi^2/3$ للتوزيع اللوجيستي بحيث إن $\pi \approx 22/7$ وبالتالي إذا ضربنا معامل البروبيت في 1.81 (تقريباً $= \pi/\sqrt{3}$) سنحصل تقريباً على معامل اللوجيت. في مثالنا الحالي، معامل انحدار الـ GPA في نموذج البروبيت هو 1.6258 بضرب ذلك في 1.81 نحصل على 2.94 وهي قيمة قريبة إلى معامل اللوجيت.

والعكس صحيح، إذا ضربنا معامل اللوجيت في 0.55 (1/1.81 =) سنحصل على معامل البروبيت. Ameniya تقترح ضرب تقدير اللوجيت في 0.625 للحصول على تقدير أفضل لتقدير معامل الانحدار باستخدام نموذج البروبيت(34).

وبالعكس، ضرب معامل انحدار نموذج البروبيت في 1.6 (1/0.625=) يعطي تقديراً جيداً لمعامل الانحدار باستخدام نموذج اللوجيت.

وبالمثل أوضحت أميميا (Amemiya) أن معاملات LPM واللوجيت مرتبطة كالتالي:

باستثناء الجزء المقطوع من المحور الصادي $\beta_{\text{LPM}}= 0.25\ \beta_{\text{logit}}$

و

الجزء المقطوع من المحور الصادي $\beta_{\text{LPM}}= 0.25\ \beta_{\text{logit}} + 0.5$

سنترك للقارئ التأكد من صحة هذه العلاقات السابقة عند تطبيقها على مثالنا الخاص بالدرجة النهائية.

11.15 نموذج التوبيت : THE TOBIT MODEL

نموذج التوبيت يعتبر امتداداً لنموذج البروبيت، وهو يعود إلى عالم الاقتصاد الحاصل على جائزة نوبل چيمس توبين James Tobin. لشرح هذا النموذج دعنا نستكمل مع مثال امتلاك المنزل السكني. في نموذج البروبيت كنا نريد تقدير احتمال امتلاك المنزل السكني كدالة في بعض المتغيرات الاجتماعية الاقتصادية. في نموذج التوبيت نهتم بإيجاد كمية الأموال التي يصرفها الفرد أو الأسرة لشراء المنزل السكني، وعلاقتها بعدد من المتغيرات الاقتصادية والاجتماعية.

وبذلك تكون لدينا مشكلة هنا ألا وهي: إذا لم يقم الفرد بشراء المنزل لاتكون لدينا أي بيانات بالنسبة له فالبيانات متاحة فقط عن المستهلكين الذين بالفعل قاموا بشراء المنزل السكني.

وبالتالي فالمستهلكون مقسمون إلى مجموعتين، الأولى مكونة من مثلاً n_1 مستهلك والمتاح عنهم بيانات عن المتغيرات المنحدرة (مثل الدخل، نسبة الفائدة في القرض، عدد أفراد الأسرة وهكذا) بالإضافة إلى بيانات عن المتغير المنحدر عليه (كمية الأموال المصروفة على شراء المنزل) أما المجموعة الثانية مكونة من n_2 مستهلك

(34) T.Amemiya, "Qualitative response model: A survey, journal of economic literative, vol. 19. 1981, pp. 481-536.

والمتاح عنهم بيانات عن المتغيرات المنحدرة فقط، ولاتوجد أي بيانات عن المتغير المنحدر عليه. العينة التي يكون متاح فيها بيانات عن المتغير المنحدر عليه لبعض المشاهدات فقط وليس كل المشاهدات تسمى عينة مراقبة(35). وبالتالي فإن نموذج التوبيت يسمى أيضاً بنموذج الانحدار المراقب. ونرى بعض الكتاب يطلقون على مثل هذه النماذج نماذج الانحدار ذات المتغير التابع المحدود، وذلك بسبب القيود المفروضة على القيم التي يأخذها المتغير المنحدر عليه.

إحصائياً فإن نموذج التوبيت يأخذ الشكل التالي:

بخلاف ذلك:

$$\begin{aligned} Y_i &= \beta_1 + \beta_2 X_i + u_i && \text{if RHS} > 0 \\ &= 0 && \text{otherwise} \end{aligned} \qquad (1.11.15)$$

بحيث إن RHS هي الجانب الأيمن من المعادلة. لاحظ أن أي متغيرات مفسرة X من الممكن إضافتها بسهولة للنموذج.

هل من الممكن تقدير انحدار (1.11.15) باستخدام n_1 مشاهدة فقط، وبدون استخدام باقي المشاهدات n_2؟ الإجابة: لا فبالنسبة لتقديرات الـ OLS للمعالم المجهولة ستكون متحيزة وغير متسقة إذا اعتمدت على الجزء n_1 فقط من المشاهدات، وهنا تكون مقدرات OLS مقدرات متحيزة حتى تقاربياً(36).

لاستعراض ذلك، دعنا ننظر إلى شكل (7.15). من الشكل نرى أنه إذا كانت Y غير مشاهدة [بسبب المراقبة (القيد المفروض على البيانات)] فإن مثل هذه المشاهدات (تساوي n_2) والتي تأخذ الرمز (X) ستقع جميعها على المحور الأفقي. إذا كانت الـ Y مشاهدة (وذلك يساوي n_1) والتي تأخذ الرمز (•) ستقع في داخل المحور السيني والصادي. وبالتالي، فإنه يتضح إذا قمنا بتقدير خط الانحدار معتمدين فقط على n_1

(35) العينة المراقبة تختلف عن العينة المبتورة والتي تكون فيها المعلومات عن المتغير المنحدر عليه متاح فقط عندما يشاهد المتغير. لن نتطرق إلى هذا الجانب هنا ولكن من يرغب في معرفة المزيد عن العينات المبتورة ممكن أن يقرأ في:

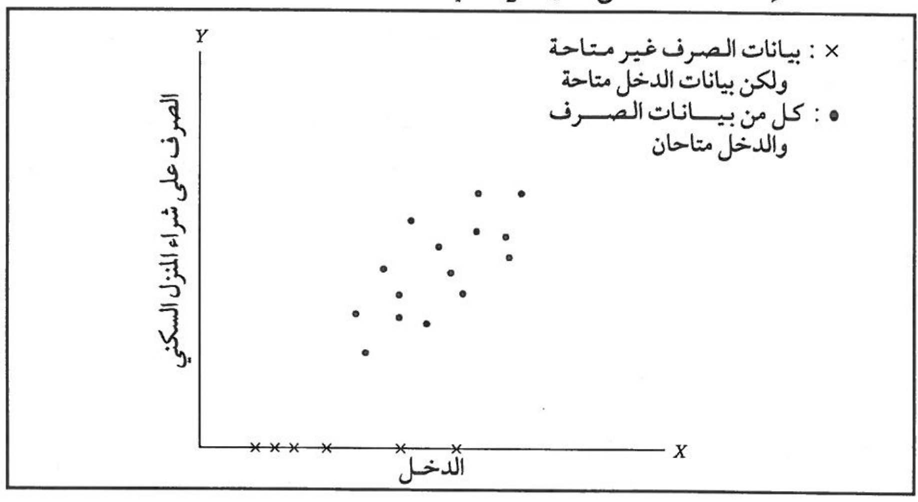
William H.Green, Econometric analysis, Prentice hall, 4th ed., Englewood Cliffs, N.J., chap. 19.

ولقراءات عميقة انظر في:

Peter Kennedy, A guide to econometrics, The MTT Press, Cambridge, Mass., 4th ed., 1988. chap. 16.

(36) التحيز يأتي من أن استخدامنا لـ n_1 مشاهدة فقط وحذف باقي المشاهدات يجعلنا لانضمن أن يكون $E(u_i)$ بالضرورة صفر. وبدون $E(u_i)=0$ فلا نستطيع ضمان أن تكون مقدرات OLS غير متحيزة. هذا التحيز من الممكن فهمه أكثر بالرجوع إلى ملحق App. 3A معادلات (4) و (5).

مشاهدة فإن الجزء المقطوع من المحور الصادي ومعاملات الانحدار (الميل) ستكون مختلفة تماماً إذا استخدمنا كل العينة $(n_1 + n_2)$ مشاهدة.



شكل (7.15) شكل يوضح كميات الأموال المصروفة على شراء المنزل السكني وعلاقتها بالدخل

كيف إذاً من الممكن تقدير نماذج انحدار توبيت (المراقبة) والتي تأخذ الشكل (1.11.15)؟ الطريقة الفعلية تعتمد على طريقة الإمكان الأعظم، والتي تفصيلها يقع خارج نطاق هذا الكتاب. ولكن القارئ من الممكن أن يعرف المزيد عن طريق ML من المراجع[37].

چيمس هيكمان James Heckman اقترح بديلاً آخر بخلاف طريقة ML والذي يعتبر أبسط نسبياً[38]. هذا البديل يشتمل على طريقة تقدير ذات خطوتين رئيسيتين. في الخطوة الأولى، نقدر أولاً احتمال أن يشتري المستهلك المنزل السكني، كما يحدث في نموذج البروبيت. وفي الخطوة الثانية نقدر نموذج (1.11.15) بإضافة متغير (يسمى مقلوب نسبة ميلز أو معدل المخاطرة) ويأخذ هذا المتغير من تقدير البروبيت. للطريقة الفعلية انظر مقالة هيكمان. طريقة هيكمان تجعلنا نحصل على مقدرات متسقة للمعالم الموجودة في (1.11.15) ولكنها تقديرات ليست لها نفس الكفاية التي تحظى بها مقدرات ML.

وبما أن معظم حزم برامج التحليل الإحصائي تحتوي على أكواد لحساب ML قد يكون استخدام هذه الحزم أفضل من استخدام طريقة هيكمان ذات الخطوتين السابق ذكرهما.

(37) انظر: - Greene, op.cit. A some what less technical discussion can be found in Richard Breen, regression models: censored, sample selected or truncated data, sage publications, new bury park, California, 1996.

(38) J.J. Heckman, "sample selection bias as a specification error", econometrica, vol. 47, pp 135- 161

مثال توضيحي لنموذج التوبيت.. نموذج راي فير للعلاقات الجنسية خارج إطار الزواج:

Illustration of the Tobit model: Ray Fair's model of extramarital affairs[39]

في مقالة شيقة وخلاقة نظرياً، جمع راي فير عينة من 601 رجل وسيدة متزوجين للمرة الأولى، وتم دراسة وتحليل إجاباتهم عن سؤال خاص بعلاقاتهم الجنسية خارج إطار الزواج[40]. المتغيرات التي تم استخدامها في الدراسة هي كالتالي:

Y = عدد مرات الممارسات الجنسية خارج الزواج في السنة الماضية 0، 1، 2، 3، 4–10 (تأخذ الكود 7)

Z_1 = 0 للسيدات و 1 للرجال.

Z_2 = عدد سنوات الزواج.

Z_4 = وجود أطفال: 0 لايوجد أطفال و 1 يوجد أطفال.

Z_5 = مستوى التدين بمقياس من 1 إلى 5 حيث 1 غير متدين على الإطلاق.

Z_6 = المستوى التعليمي ويحسب بعدد السنوات: 9= تعليم أساسي، 12= تعليم ثانوي، دكتوراه وخلافه= 20.

Z_7 = الوظيفة، مقياس holling shead من 7–1.

Z_8 = تقييم شخصي للزواج، 1= غير سعيد على الإطلاق، 5= سعيد جداً.

من 601 مفردة يوجد 451 مفردة بدون أي علاقات جنسية خارج الزواج و 150 مفردة لها تجربة أو أكثر في ذلك.

بنفس فكرة شكل (7.15)، إذا رسمنا عدد هذه التجارب الجنسية على المحور الرأسي، وعلى سبيل المثال، التعليم على المحور الأفقي سيكون لدينا 451 مفردة تقع على المحور الأفقي مباشرة، وبالتالي لدينا عينة مراقبة ونموذج التوبيت يكون ملائماً للبيانات في مثل هذه الحالة.

جدول (16.15) يعطي تقديرات مختلفة للنموذج باستخدام كل من OLS (غير المناسبة) وطريقة ML (المناسبة). وكما نرى فإن OLS تشتمل على 451 مفردة التي لم يكن لها تجارب جنسية خارج الزواج و 150 مفردة لها تجربة أو أكثر. طريقة ML تأخذ ذلك في الاعتبار بشكل واضح بخلاف طريقة الـ OLS وهذا هو الفرق بين هاتين

---

(39) Ray fair, "A theory of extramarital affairs", Journal of Political Economy, vol. 86, 1978, pp. 45-61. For the article and the data, see http://fairmodel.econ,yale.edu/rayfairo/pdf/1978DAT.ZIP.

(40) في عام 1969 نشرت مجلة علم النفس اليوم استمارة بحثية مكونة من 101 سؤال عن الجنس وطلبت من القراء إرسال إجاباتهم بالبريد. في عدد يوليو 1970 نتائج البحث تم استعراضها وكان هناك حوالي 2000 رد تم تجميعها إلكترونياً. راي فير استخرج من هذه الردود عينة من 601 مفردة.

الطريقتين. لأسباب سبق شرحها من الأفضل استخدام تقديرات ML عن تقديرات OLS. المعاملات من النموذجين من الممكن تفسيرها كما سبق، وفعلنا مع أي معاملات انحدار أخرى. المعامل ذو الإشارة السالبة للمتغير Z_8 (السعادة الزوجية) يعني أنه كلما زادت السعادة الزوجية كلما قل مؤشر التجارب الجنسية خارج إطار الزواج وهذه نتيجة متوقعة.

كما سبق وذكرنا، إذا كان اهتمامنا باحتمال التجارب الجنسية خارج إطار الزواج وليس عدد مثل هذه التجارب فمن الممكن استخدام نموذج البروبيت.

جدول (16.15) تقديرات OLS والتوبيت للتجارب الجنسية خارج إطار الزواج

Explanatory variable	OLS estimate	Tobit estimate
Intercept	5.8720 (5.1622)*	7.6084 (1.9479)†
Z_1	0.0540 (0.1799)	0.9457 (0.8898)
Z_2	−0.0509 (−2.2536)	−0.1926 (−2.3799)
Z_3	0.1694 (4.1109)	0.5331 (3.6368)
Z_4	−0.1426 (−0.4072)	1.0191 (0.7965)
Z_5	−0.4776 (−4.2747)	−1.6990 (−4.1906)
Z_6	−0.0137 (−0.2143)	0.0253 (0.1113)
Z_7	0.1049 (1.1803)	0.2129 (0.6631)
Z_8	−0.7118 (−5.9319)	−2.2732 (−5.4724)
R^2	0.1317	0.1515

* : الأرقام داخل الأقواس هي قيم الـ t.

T : الأرقام داخل الأقواس هي قيم التوزيع الطبيعي القياسي Z.

لاحظ أن: من العدد الكلي للمفردات 601 هناك 451 مفردة تأخذ 0 للمتغير التابع (عدد التجارب الجنسية خارج إطار الزواج) و 150 مفردة لها قيم مختلفة عن الصفر.

جدول (17.15) المتغير التابع : y star

الطريقة : البروبيت الثنائي – ML

العينة : 1–601

المفردات المستخدمة : 601

النتائج تم الحصول عليها بعد 5 تكرارات

Variable	Coefficient	Std. error	Z statistic	Probability
C	0.779402	0.512549	1.520638	0.1284
Z_1	0.173457	0.137991	1.257015	0.2087
Z_2	-0.024584	0.010418	-2.359844	0.0183
Z_3	0.054343	0.018809	2.889278	0.0039
Z_4	0.216644	0.165168	1.311657	0.1896
Z_5	-0.185468	0.051626	-3.592551	0.0003
Z_6	0.011262	0.029517	0.381556	0.7028
Z_7	0.013669	0.041404	0.330129	0.7413
Z_8	-0.271791	0.053475	-5.082608	0.0000

Mean dependent var	0.249584	S.D. dependent var	0.433133
S.E. of regression	0.410279	Akaike info criterion	1.045584
Sum squared resid	99.65088	Schwarz criterion	1.111453
Log likelihood	-305.1980	Hannan-Quinn criter.	1.071224
Restr. log likelihood	-337.6885	Avg. log likelihood	-0.507817
LR statistic (8 df)	64.98107	McFadden R-squared	0.096215
Probability (LR stat)	4.87E-11		
Obs with Dep = 0	451	Total obs	601
Obs with Dep = 1	150		

بإعطاء المتغير Y القيمة 0 عندما تكون المشاهدة ليس لديها أي تجارب جنسية خارج إطار الزواج، و $Y=1$ عندما تكون المشاهدة لها مثل هذه التجارب يعطي النتائج الموجودة في جدول (17.15). باستخدام نموذج البروبيت من الممكن للقارئ أن يفسر بنفسه النتائج الموجودة في الجدول السابق.

12.15 نمذجة بيانات العدد.. نموذج انحدار بواسون:

THE POISSON REGRESSION MODEL.. MODELING COUNT DATA

هناك العديد من الظواهر التي يكون المتغير المنحدر عليه فيها له طبيعة العد، فمثلاً عدد الإجازات التي تأخذها أسرة ما سنوياً، عدد براءات الاختراع المسجلة تحصل عليه مؤسسة ما سنوياً، عدد مرات الذهاب إلى طبيب الأسنان أو أي طبيب آخر سنوياً، عدد مرات الذهاب إلى محل البقالة أسبوعياً، عدد المخالفات المرورية سرعة زائدة أو وضع السيارة في أماكن غير مسموح بها سنوياً، عدد أيام البقاء في مستشفى ما لتلقي العلاج خلال فترة معينة، عدد السيارات المارة في نقطة مرور ما خلال 5 دقائق مثلاً وهكذا.

المتغير محل الدراسة في كل من الحالات السابقة هو متغير متقطع يأخذ عدداً محدداً من القيم. أحياناً تكون البيانات العددية نادرة أو قليلة الحدوث مثل عدد مرات الإصابة بصعقة كهربائية في خلال أسبوع، أو الفوز بأكثر من ورقة يانصيب في خلال أسبوعين أو حدوث أكثر من أزمتين أو أكثر في القلب في خلال أربعة أسابيع. كيف يمكن نمذجة مثل هذه الظواهر؟

كما تم استخدم توزيع برنولي لنمذجة البيانات من نوع (نعم/ لا) في نماذج الاحتمال الخطية، فإن التوزيع الاحتمالي المناسب للبيانات العددية هو توزيع بواسون الاحتمالي.

دالة التوزيع الاحتمالي لتوزيع بواسون تأخذ الشكل التالي[41].

$$f(Y_i) = \frac{\mu^Y e^{-\mu}}{Y!} \qquad Y = 0, 1, 2, \ldots \qquad \textbf{(1.12.15)}$$

(41) انظر في أي كتاب أساسي في الإحصاء للحصول على مزيد من التفاصيل عن هذا التوزيع.

بحيث إن $f(Y)$ ترمز لاحتمال أن يأخذ المتغير Y قيمة غير سالبة و $Y!$ (تقرأ مضروب Y) هي عبارة عن $Y! = Y \times (Y-1) \times (Y-2) \times 2 \times 1$. من الممكن إثبات التالي:

$$E(Y) = \mu \qquad (2.12.15)$$

$$Var(Y) = \mu \qquad (3.12.15)$$

لاحظ خاصية مميزة لتوزيع بواسون: تباينه يساوي توقعه. نموذج انحدار بواسون يأخذ الشكل التالي:

$$Y_i = E(Y_i) + u_i = \mu_i + u_i \qquad (4.12.15)$$

بحيث إن Y's متغيرات مستقلة لها توزيع بواسون بتوقع μ_i لكل مفردة وهذا التوقع هو:

$$\mu_i = E(Y_i) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} \qquad (5.12.15)$$

بحيث إن X's هي بعض المتغيرات التي قد يكون لها تأثير على قيمة التوقع. على سبيل المثال، إذا كان المتغير العددي هو عدد مرات الذهاب إلى متحف مترو بوليتان الفني في نيويورك في سنة ما، هذا العدد سيعتمد على عدد من المتغيرات منها، دخل المستهلك، سعر تذكرة الدخول، المسافة إلى المتحف، وتكلفة موقف السيارة.

لأسباب تتعلق بالتقدير سنكتب النموذج كالتالي:

$$Y_i = \frac{\mu^Y e^{-\mu}}{Y!} + u_i \qquad (6.12.15)$$

بحيث إن μ تم التعويض عنها بالمعادلة (5.12.15) ويمكن ملاحظة أن نموذج الانحدار الناتج سيكون غير خطي في المعلمات، ويحتاج إلى طريق تقدير الانحدار غير الخطية التي تم مناقشتها في الفصل السابق. دعنا نستعرض مثالاً واقعياً لنرى كيفية التعامل بمثل هذا النموذج.

مثال توضيحي.. دراسة على كبار السن وتكرار الوقوع:

An illustrative example: Geriatric study on frequency of falls

البيانات المستخدمة هنا قام بتجميعها نينيري وآخرون[42]. البيانات خاصة بـ 100 فرد أعمارهم 65 سنة فأكثر. هدف الدراسة هو تسجيل عدد مرات الوقوع (Y=) لهؤلاء الأفراد وعلاقة ذلك بالنوع (X_2=0 للإناث و 1= للذكور)، مؤشر للتوازن(X_3)، مؤشر

(42) John Neter, Michael H.kutner, Christopher J Nachtshein, and William Wasserman, Applied Regression Models, Irwin, 3d ed., Chicago, 1996. The data were obtained from the data disk included in the book and refer to exercise 14.28.

للقوة (X_4). كلما زاد مؤشر القوة كلما كان ذلك دليلاً على توازن أكثر للشخص، وكلما زاد مؤشر القوة كلما كان ذلك دليلاً على قوة أكثر للشخص. لمعرفة ما إذا كانت تمارين الأيروبيكس ذات مستوى بسيط أو مستوى أكثر تقدماً لها تأثير على عدد مرات الوقوع، قام الباحث بإدخال متغير إضافي X_1 وأطلق عليه اسم متغير طارئ intervention بحيث إن $X_1=0$ إذا قام الفرد بتدريبات إيروبيكس بسيطة و $X_1=1$ إذا قام الفرد بتدريبات أكثر تقدماً. تم توزيع الأشخاص على برنامج التدريب بشكل عشوائي باستخدام eviews 4 حصلنا على النتائج الموجودة في جدول (18.15).

جدول (18.15)
المتغير التابع: Y
العينة: 100 - 1
النتائج تم الحصول عليها بعد 7 تكرارات
Y=EXP(C(0)+C(1)*X1+C(2)*X2+C(3)*X3+C(4)*X4)

	Coefficient	Std. error	t statistic	Probability
C(0)	0.37020	0.3459	1.0701	0.2873
C(1)	-1.10036	0.1705	-6.4525	0.0000
C(2)	-0.02194	0.1105	-0.1985	0.8430
C(3)	0.01066	0.0027	3.9483	0.0001
C(4)	0.00927	0.00414	2.2380	0.0275

$R^2 = 0.4857$ Adjusted $R^2 = 0.4640$
Log likelihood = -197.2096 Durbin-Watson statistic = 1.7358

لاحظ أن: ( )Exp تعني e (أساس اللوغاريتم الطبيعي) مرفوعة للآس الموجود في ( ).

تفسير النتائج:

ضع في الاعتبار ما تم الحصول عليه في جدول 18.15 هو القيمة المقدرة لتوقع المشاهدة j، $\hat{\mu}_i$ وبالتالي ماقمنا بتقديره هو:

$$\hat{\mu}_i = e^{0.3702-1.100366X_{1i}-0.02194X_{2i}+0.0106X_{3i}+0.00927X_{4i}} \qquad (7.12.15)$$

للحصول على القيمة الفعلية لمتوسط المفردة j نعوض بقيم المتغيرات X المختلفة التابعة لهذه المفردة. على سبيل المثال، المفردة رقم 99 لها هذه القيم:

$Y= 4, X_1= 0, X_2= 1, X_3= 50, X_4= 56$ وبالتعويض بهذه القيمة في معادلة (7.12.15) نحصل على $\hat{\mu}_{99}=3.3538$ كقيمة مقدرة لمتوسط المشاهدة رقم 99. القيمة الفعلية للمتغير y لهذه المفردة هي 4.

والآن إذا أردنا الحصول على احتمال أن تقع مفردة مثل المفردة رقم 99 أقل من 5 مرات في العام فيتم ذلك كالتالي:

$$P(Y<5) = P(Y=0)+P(Y=1)+P(Y=2)+P(Y=3)+P(Y=4)$$

$$= \frac{(3.3538)^0 e^{-3.3538}}{0!} + \frac{(3.3538)^1 e^{-3.3538}}{1!} + \frac{(3.3538)^2 e^{-3.3538}}{2!} + \frac{(3.3538)^3 e^{-3.3538}}{3!} + \frac{(3.3538)^4 e^{-3.3538}}{4!}$$

$$= 0.7491$$

من الممكن أيضاً الحصول على التأثير الحدي أو الجزئي لمتغير منحدر ما على متوسط المتغير Y كالتالي: في مثالنا التوضيحي الحالي، دعنا نفترض أننا نريد معرفة تأثير زيادة مؤشر القوة X_4 بوحدة واحدة على المتوسط Y بما أن:

$$\mu = e^{C_0+C_1X_{1i}+C_2X_{2i}+C_3X_{3i}+C_4X_{4i}} \quad (8.12.15)$$

نريد معرفة $\partial\mu/\partial X_4$. باستخدام قاعدة السلسلة الحسابية، من الممكن بسهولة إثبات أن ذلك يساوي

$$\frac{\partial\mu}{\partial X_4} = C_4 e^{C_0+C_1X_{1i}+C_2X_{2i}+C_3X_{3i}+C_4X_{4i}} = C_4\mu \quad (9.12.15)$$

وهذا يعني أن معدل التغير في القيمة المتوسطة (التوقع) بالنسبة لمتغير منحدر ما يساوي معامل هذا المتغير مضروباً في القيمة المتوسطة. بالطبع القيمة المتوسطة μ تعتمد على كل القيم التي تأخذها كل المتغيرات المنحدرة في النموذج. وهذا يشبه نموذج اللوجيت، ونموذج البروبيت السابق ذكرهما من قبل، عندما كان التأثير الحدي لمتغير ما يعتمد أيضاً على القيم التي تأخذها كل المتغيرات الموجودة في النموذج.

وبالعودة إلى **المعنوية الإحصائية الخاصة بمعاملات الانحدار الفردية**، نجد أن الجزء المقطوع من المحور الصادي، ومعامل المتغير X_2 لكل منهما على حدى معنوية إحصائية. ولكن لاحظ أن الإخطاء القياسية المعطاة في الجدول هي تقاربية، وبالتالي قيم t لابد من تفسيرها تقاربياً إيضاً. كما سبق وذكرنا بوجه عام نتائج التقدير غير الخطي بطريقة التكرار يمكن الاعتماد عليها في إطار العينات كبيرة الحجم فقط.

كنتيجة نهائية لكل ما سبق عرضه عن نموذج انحدار بواسون نجد أنه يفترض عدة افتراضات مقيدة كأن يكون التوقع والتباين الخاص بعملية بواسون متساويين، ويكون احتمال وقوع أي حدث ثابت عند أي نقطة زمنية.

13.15 موضوعات أخرى في نماذج الانحدارات ذات الاستجابة النوعية FURTHER TOPICS IN QUALITATIVE RESPONSE

كما ذكرنا في البداية، فإن موضوع نماذج الانحدار ذات الاستجابة النوعية يعتبر موضوعاً ضخماً. ما استعرضناه في هذا الفصل، بعض من النماذج الأساسية في إطار هذا الموضوع. لمن يرغب في مزيد من التفاصيل عن هذا الموضوع، سنستعرض باختصار شديد بعض النماذج الأخرى في هذا الإطار. لن نتطرق لهذا الموضوع بالتفصيل حيث يقع ذلك خارج نطاق هذا الكتاب.

نماذج اللوجيت والبروبيت الترتيبية : Ordinal logit and probit models

في نماذج اللوجيت والبروبيت الثنائية كنا مهتمين بدراسة المتغير التابع الذي له طبيعة (نعم/ لا) ولكن غالباً ما يكون المتغير التابع أو المنحدر عليه له أكثر من نتيجتين، وغالباً ما تكون هذه النتائج ترتيبية بطبيعتها، وبالتالي لا يمكن التعبير عنها في مقياس فئوي. عادة في الأبحاث ذات الاستمارات البحثية غالباً ما تكون الإجابات تأخذ مقياسًا من نوع ليكرت مثل "أوافق بشدة" و "موافق نوعاً ما" أو "غير موافق بشدة" أو مثلاً الإجابات في الأبحاث الخاصة بالتعليم قد تكون "أقل من المرحلة الثانوية"، "المرحلة الثانوية"، "المرحلة الجامعية" أو "مرحلة الدراسات العليا". غالباً ما تأخذ هذه الإجابات أكواداً مثل 0 (أقل من المرحلة الثانوية)، 1 (المرحلة الثانوية)، 2 (الجامعة)، و 3 (دراسات عليا). هذا هو مقياس ترتيبي، بحيث إن هناك ترتيبًا واضحاً بين الفئات (الطبقات) المختلفة، ولكن لا تستطيع القول بأن 2 (التعليم الجامعي) هي ضعف 1 (التعليم الثانوي) أو لا تستطيع القول بأن 3 (الدراسات العليا) هي ثلاث مرات قدر 1 (التعليم الثانوي).

لدراسة ظاهرة مثل ذلك، من الممكن عمل امتداد لنماذج اللوجيت والبروبيت الثنائية، بحيث تأخذ في الاعتبار الطبقات المتعددة المرتبة. سنحتاج إلى خلفية رياضية جيدة نوعاً ما، حيث إننا سنحتاج لاستخدام توزيع طبيعي متعدد وتوزيع لوجيتيكي متعدد ليسمح بإدخال الطبقات المتعددة المرتبة. لمعرفة المزيد عن الخلفية الرياضية لهذا الموضوع وتطبيقاته من الممكن الاستعانة بكتاب جرين ومادالا السابق ذكره من قبل. ويمكن للقارئ الاستعانة بكتاب آخر مماثل وهو ليو مانوجراف(43). حزم البرامج الإلكترونية مثل shazan و eviews و limdep لديها أكواد خاصة لتقدير نماذج اللوجيت والبروبيت الترتيبية.

(43) Tim Futing Liao, op cit.

نماذج اللوجيت والبروبيت الإسمية المتعددة :
Multinomial logit and probit models

في نماذج اللوجيت والبروبيت المرتبة، يكون المتغير المستجيب له أكثر من طبقتين مرتبتين، ولكن هناك حالات يكون المتغير المنحدر عليه غير ترتيبي. فعلى سبيل المثال، وسيلة الانتقال إلى العمل قد تكون دراجة، دراجة بخارية، سيارة، أتوبيس أو قطار. فنجد أن المتغير التابع هو متغير طبقي، ولكن بدون ترتيب داخل هذه الطبقات، فهي طبقات إسمية بطبيعتها. مثال آخر التقسيم المهني مثل غير ماهر، متوسط المهارة، عالي المهارة. مرة أخرى لايوجد ترتيب. بالمثل أنواع المهن المختلفة مثل مهن حرة، قطاع خاص، العمل في القطاع العام (المحلي) أو العمل في الحكومة الاتحادية يعتبر متغيراً اسمياً بطبيعته.

أساليب نماذج اللوجيت والبروبيت الإسمية المتعددة تكون مناسبة لدراسة هذه المتغيرات ذات الطبقات الإسمية. مرة أخرى، نحتاج إلى خلفية رياضية خاصة لدراسة هذا الموضوع. المرجع السابق ذكره يمكن استخدامه لفهم هذه الأساليب. ومن الممكن أيضاً استخدام حزم البرامج الإلكترونية السابق ذكرها لتقدير هذه النماذج عندما تحتاج المشكلة محل الدراسة لذلك.

نماذج البقاء : Duration models

دعنا نستعرض الأسئلة التالية: (1) ما هي العوامل المحددة لمدة البطالة؟ (2) ما هي العوامل المحددة لعمر لمبة الإنارة؟ (3) ما هي العوامل المحددة لمدة طلعة جوية ما (غارة)؟ (4) ما هي العوامل المحددة لفترة بقاء مريض إيجابي HIV على قيد الحياة؟

المواضيع المتعلقة بنماذج البقاء غالباً ما تعرف باسم تحليل الدوام (البقاء) أو تحليل البيانات (الوقت حتى حدث ما). في كل الأمثلة السابق ذكرها المتغير الأساسي هو الوقت اللازم لحدوث شيء ما، والذي يعتبر متغيراً عشوائياً مرة أخرى الخلفية الرياضية للموضوع تشتمل على CDFs و PDFxs الخاصة بنماذج الاحتمال المختلفة. على الرغم من صعوبة تفاصيل هذه الطريقة، فإن هناك بعض الكتب التي يمكن فهمها بسهولة نوعاً ما عن هذا الموضوع[44]. الحزم الإحصائية مثل Stat وLimdep ممكن استخدامها بسهولة لتقدير نماذج البقاء. هذه الحزم تعطي أمثلة توضيحية لكي ترشد الباحث عن كيفية التعامل مع مثل هذه النماذج.

(44) انظر على سبيل المثال: David W. Hosmer, Jr., and Stanley Lemeshow, Applied Survival Analysis, John Wiley & Sons, New York. 1999.

14.15 التلخيص والنتائج : SUMMARY AND CONCLUSIONS

1 - نماذج الانحدار ذات الاستجابة النوعية يكون فيها المتغير التابع أو المنحدر عليه ليس متغيراً كمياً وليس له مقياس فئوي.

2 - أبسط نماذج الانحدار ذات الاستجابة النوعية المتاحة هو النموذج الثنائي عندما يكون المتغير المنحدر عليه من نوع (نعم/ لا) أو (حضور/ عدم حضور) وهكذا.

3 - أبسط نماذج الانحدار الثنائية المتاحة هو نموذج الاحتمال الخطي (LPM) عندما ينحدر المتغير المستجيب الثنائي على عدد من المتغيرات المفسرة باستخدام طريقة OLS التقليدية. وعموماً الموضوع ليس بسيطاً كما يبدو، حيث إن LPM تعاني من العديد من مشاكل التقدير حتى لو استطعنا التغلب على بعض من هذه المشاكل، فإن نقطة الضعف الرئيسية في الـ LPM هي أنه يفترض أن احتمال حدوث شيء ما يزداد خطياً مع مستوى التغير المنحدر. وهذا الفرض المقيد من الممكن الاستغناء عنه إذا استخدمنا نماذج اللوجيت والبروبيت.

4 - في نموذج اللوجيت يكون المتغير التابع هو لوغاريتم نسبة الأوزان، والذي هو بدوره دالة خطية في المتغيرات المنحدرة. دالة الاحتمال المستخدمة في نموذج اللوجيت هي دالة التوزيع اللوجيتيكي. إذا كانت البيانات المتاحة موجودة في صورة تجميعية (مجموعات) فإننا نستطيع استخدام OLS لتقدير نموذج اللوجيت، ولكن يجب أن نأخذ في الاعتبار مشكلة اختلاف التباين الموجودة في مقدار الخطأ. أما إذا كانت البيانات المتاحة موجودة بصورة منفردة أو جزئية لابد من استخدام طرق التقدير غير الخطية في المعلمات.

5 - إذا اخترنا التوزيع الطبيعي على أنه التوزيع الاحتمالي المناسب، فمن الممكن استخدام نموذج البروبيت. هذا النموذج يعتبر صعباً رياضياً نوعاً ما، حيث إنه يشتمل على بعض التكاملات. ولكن لكل الأغراض العملية المختلفة، فإن نماذج اللوجيت والبروبيت تعطي نتائج متماثلة. في الواقع الاختيار بينهما يعتمد على سهولة الحساب، والذي لايعتبر مشكلة حقيقية الآن طالما هناك توفر للعديد من الحزم الإحصائية المتخصصة.

6 - إذا كان المتغير المستجيب له طبيعة عددية، فإن النموذج المستخدم عادة هو نموذج انحدار بواسون والذي يعتمد على دالة احتمال توزيع بواسون.

7 - نموذج التوبيت هو نموذج قريب من نموذج البروبيت، وهو معروف أيضاً باسم نموذج الانحدار المراقب. في هذا النموذج، المتغير التابع يشاهد فقط إذا تحققت

بعض الشروط (أو قد يكون شرطًا واحدًا). وبالتالي فإن السؤال الخاص بكمية الأموال الممكن صرفها لشراء سيارة ما يصبح ذا معنى فقط عندما يبدأ الفرد بشراء سيارة ما. عموماً مادالا Maddala لاحظ أن نموذج توبيت "من الممكن استخدامه فقط في الحالات التي يكون فيها المتغير الكامن (بمعنى المتغير الأساسي للظاهرة) من الممكن في الأصل أن يأخذ قيماً سالبة، وتكون قيم الصفر المشاهدة (المسجلة) هي تابعة للقيد وغير مشاهدة"(45).

8 - هناك بعض النماذج التي تعتبر امتداداً لنموذج الانحدار ذي الاستجابة الثنائية. هذه النماذج تشتمل على نموذج البروبيت و اللوجيت الترتيبية، ونماذج اللوجيت والبروبيت الإسمية المتعددة. المنطق وراء هذه النماذج جميعاً يعتبر منطقاً واحداً مماثلاً لنماذج اللوجيت والبروبيت الأبسط، على الرغم من أن الجانب الرياضي يصبح أكثر تعقيداً.

9 - أخيراً، استعرضنا باختصار ما يسمى نماذج البقاء التي يدرس فيها الفترة الزمنية لظاهرة ما مثل البطالة أو الإصابة بمرض ما إلى ما غير ذلك من هذه الظواهر التي تعتمد على عوامل عديدة مختلفة. في مثل هذه النماذج الفترة الزمنية وفترة البقاء تصبح المتغير المراد دراسته وتفسيره.

تمارين: EXERCISES

أسئلة: Questions

1.15 بالعودة إلى البيانات المعطاة في جدول (2.15). إذا كانت $\hat{Y}_i$ سالبة، افترض أنها تساوي 0.01 وإذا كانت أكبر من الواحد الصحيح، افترض أنها تساوي 0.99. أعد حساب الأوزان W_i ومقدرات الـ LPM باستخدام WLS. قارن نتائجك مع النتائج المعطاة في (11.2.15) وعلق على ذلك.

2.15 باستخدام بيانات امتلاك المنزل السكني الموجودة في جدول (1.15)، تقديرات طريقة الإمكان الأعظم لنموذج اللوجيت هي كالتالي:

$$\hat{L}_i = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right) = -493.54 + \quad 32.96 \text{ income}$$
$$t = \quad (-0.000008)(0.000008)$$

علق على هذه النتائج، مع الوضع في الاعتبار أن كل بيانات الدخل الأكثر من 16 (ألف دولار) ممثلة في Y=1 وكل قيم الدخل الأخرى الأقل من 16 ممثلة في Y=0. ماذا تتوقع في مثل هذه الحالة؟

(45) G.S. Maddala, Introduction of econometrics, 2d ed., Macmillan, New York, 1992, p. 342

3.15 في دراسة عن شراء السلع الاستهلاكية Y (Y=1 في حالة الشراء، Y=0 في حالة عدم الشراء) كدالة في عدد 5 متغيرات. تمت الدراسة على 762 أسرة حصل Janet A.Fisher* (جانيت فيشر) على النتائج التالية:

Explanatory variable	Coefficient	Standard error
Constant	0.1411	—
1957 disposable income, X_1	0.0251	0.0118
(Disposable income = X_1)^2, X_2	−0.0004	0.0004
Checking accounts, X_3	−0.0051	0.0108
Savings accounts, X_4	0.0013	0.0047
U.S. Savings Bonds, X_5	−0.0079	0.0067
Housing status: rent, X_6	−0.0469	0.0937
Housing status: own, X_7	0.0136	0.0712
Monthly rent, X_8	−0.7540	1.0983
Monthly mortgage payments, X_9	−0.9809	0.5162
Personal noninstallment debt, X_{10}	−0.0367	0.0326
Age, X_{11}	0.0046	0.0084
Age squared, X_{12}	−0.0001	0.0001
Marital status, X_{13} (1 = married)	0.1760	0.0501
Number of children, X_{14}	0.0398	0.0358
(Number of children = X_{14})^2, X_{15}	−0.0036	0.0072
Purchase plans, X_{16} (1 = planned; 0 otherwise)	0.1760	0.0384
$R^2 = 0.1336$		

لاحظ أن: كل المتغيرات الاقتصادية مقاسة بالألف دولار.
حالة المنزل: الإيجار (1 إذا كان المنزل مؤجراً، 5 بخلاف ذلك)
حالة المنزل: الملكية (1 إذا كان المنزل مملوكاً، 5 بخلاف ذلك)

المصـــدر: Janet A.Fisher, "An analysis of consumer good expenditure", The review of economics and statistics, vol. 64, no.1, table, 1962, p.27

(a) علق بوجه عام على تقدير النموذج من خلال المعادلة.

(b) كيف يمكن تفسير معامل 0.0051− الانحدار المرتبط بمتغير الحساب الجاري؟ كيف يمكن التعليق بشكل مقبول على الإشارة السالبة للمتغير؟

(c) ما هو المنطق وراء استخدام مربع العمر كمتغير، ومربع عدد الأطفال كمتغير؟ لماذا نجد الإشارة سالبة في كل من الحالتين؟

(d) بافتراض تساوي جميع المتغيرات بالصفر باستثناء متغير الدخل. أوجد الاحتمال الشرطي لشراء سلعة استهلاكية بافتراض أن دخل هذه الأسرة $20.000 .

(e) قدر الاحتمال الشرطي لامتلاك سلعة أو سلع استهلاكية بشرط:

X_1= \$15.000, X_3= \$3000, X_4= \$5000, X_6= 0, X_8= \$500, X_9= \$300, X_{10}= \$0, X_{11}= 35, X_{13}= \$1, X_{14}= 2, X_{16}= 0

4.15 قيمة R^2 في جدول (3.15) الخاص بمثال نموذج الانحدار للمشاركة في قوة العمل هي 0.175 وهي تعتبر قيمة صغيرة نسبياً. هل من الممكن اختبار معنوية هذه القيمة؟ ما هو الاختيار المستخدم ولماذا؟ علق بوجه عام على قيمة R^2 في النماذج الأخرى المماثلة.

5.15 قدر احتمال امتلاك المنزل السكني وفقاً لمستويات مختلفة من الدخل في الانحدار الموجود في (1.7.15) وضح في شكل بياني علاقة هذه الاحتمالات مع الدخل، وعلق على هذه النتيجة.

6.15(*) في نموذج بروبيت المعطى في جدول (11.15) اثبت أن الجزء المقطوع من المحور الصادي يساوي $\frac{-\mu_x}{\sigma_x}$ والميل يساوي $\frac{1}{\sigma_x}$ بحيث إن μ_x و σ_x هما التوقع والانحراف المعياري لـ X.

7.15 من بيانات تم تجميعها من 54 عاصمة في العالم من مناطق إحصائية قياسية، Demairs ديمارز قدر نموذج اللوجيت التالي لتفسير معدلات الجريمة المرتفعة في مقابل معدلات الجريمة المنخفضة (+) وكان النموذج المقدر كالتالي:

$$\ln\hat{O}_i = 1.1387 + 0.0014P_i + 0.0561C_i - 0.4050R_i$$
$$\text{Se} = \quad (0.0009) \quad (0.0227) \quad (0.1568)$$

بحيث إن O = أوزان معدل الجريمة المرتفعة، P= 1980 حجم المجتمع بالآلاف، C = معدل نمو السكان من 1970 إلى 1980، R = حاصل القراءة ويرمز Se إلى الأخطاء القياسية التقاربية.

(a) كيف يمكنك تفسير المعاملات المختلفة؟

(b) أي من المعاملات يعتبر له معنوية إحصائية فردية؟

(c) ما أثر الزيادة بمقدار الوحدة في متغير حاصل القراءة R على أوزان معدل الجريمة المرتفع؟

(d) ما أثر الزيادة بنسبة واحدة في معدل نمو السكان على أوزان معدل الجريمة المرتفع؟

8.15 قارن وعلق على انحدار OLS و WLS في (3.7.15) و (1.7.15).

(*) اختياري

(+) Demairs, op. cit., p. 46.

مسائل : Problems

9.15 من بيانات مسح ميزانية الأسرة في عام 1980، والخاص بالمركز الرئيسي الهولندي للإحصاء، J.S. Cramer كرامير حصل على نموذج اللوجيت التالي، والمعتمد على عينة حجمها 2380 أسرة (النتائج المعطاة هنا تم الحصول عليها بطريقة الإمكان الأعظم بعد التكرار الثالث)(*). الهدف من نموذج اللوجيت هو تقدير امتلاك السيارة كدالة في لوغاريتم الدخل، امتلاك السيارة هو متغير ثنائي: $Y=1$ إذا كانت الأسرة تمتلك السيارة وصفر بخلاف ذلك.

$$\hat{L}_i = -2.77231 + 0.347582 \text{ ln income}$$
$$t = \quad (-3.35) \qquad (4.05)$$
$$X^2(1 \text{ df}) = 16.681 \ (p \text{ value} = 0.0000)$$

بحيث إن $\hat{L}_i$ تقدير اللوجيت و ln income هو لوغاريتم الدخل. قيمة X^2 تقيس جودة التوفيق للنموذج.

(a) فسر نموذج اللوجيت المقدر.

(b) من نموذج اللوجيت المقدر، كيف يمكنك الحصول على شكل احتمال امتلاك السيارة؟

(c) ما هو احتمال أن تملك الأسرة السيارة ولها دخل 20.000؟ وعند مستوى دخل 25.000؟ ما هو معدل التغير في الاحتمال عند مستوى دخل 20.000؟

(d) علق على المعنوية الإحصائية لنموذج اللوجيت المقدر.

10.15 احصل على المعادلة (8.2.15).

11.15 في دراسة مهمة عن معدلات التخرج الجامعي للمسجلين بالجامعة من كل المدارس الثانوية والمسجلين بالجامعة من الزنوج فقط، Bowen and Bok بون وبوك حصلا على النتائج التالية الموجودة في جدول (19.15) والخاصة بنموذج اللوجيت(**).

(*) J.S. Cramer, An introduction to the logit model for economist, 2d ed. Published and distributed is Timberlake consultans. Ltd., 2001, p. 33.

النتائج الإحصائية تم الحصول عليها باستخدام الحزمة الإحصائية pc give 10 المنشورة عن طريق Timberlake الاستشاري صفحة 51 .

(**) William G. Bowen and Derek Bok, The Shape of the River: Long Term Consequences of Considering Race in Collega and University Admissions, Princeton University Press, Princeton, N. J., 1998, p. 381.

جدول (19.15) نموذج انحدار لوجيتيكي للتنبؤ بمعدلات التخرج، 1989 (entering cohort) ؟

Variable	All matriculants			Black only		
	Parameter estimate	Standard error	Odds ratio	Parameter estimate	Standard error	Odds ratio
Intercept	0.957	0.052	—	0.455	0.112	—
Female	**0.280**	0.031	1.323	**0.265**	0.101	1.303
Black	**−0.513**	0.056	0.599			
Hispanic	**−0.350**	0.080	0.705			
Asian	**0.122**	0.055	1.130			
Other race	**−0.330**	0.104	0.719			
SAT > 1299	**0.331**	0.059	1.393	0.128	0.248	1.137
SAT 1200–1299	**0.253**	0.055	1.288	0.232	0.179	1.261
SAT 1100–1199	**0.350**	0.053	1.420	0.308	0.149	1.361
SAT 1000–1099	**0.192**	0.054	1.211	0.141	0.136	1.151
SAT not available	**−0.330**	0.127	0.719	0.048	0.349	1.050
Top 10% of high school class	**0.342**	0.036	1.407	**0.315**	0.117	1.370
High school class rank not available	−0.065	0.046	0.937	−0.065	0.148	0.937
High socioeconomic status (SES)	**0.283**	0.036	1.327	**0.557**	0.175	1.746
Low SES	**−0.385**	0.079	0.680	**−0.305**	0.143	0.737
SES not available	**0.110**	0.050	1.116	0.031	0.172	1.031
SEL-1	**1.092**	0.058	2.979	**0.712**	0.161	2.038
SEL-2	**0.193**	0.036	1.212	**0.280**	0.119	1.323
Women's college	**−0.299**	0.069	0.742	0.158	0.269	1.171
Number of observations	32,524			2,354		
−2 log likelihood						
Restricted	31,553			2,667		
Unrestricted	30,160			2,569		
Chi square	1,393 with 18 d.f.			98 with 14 d.f.		

لاحظ أن: المعاملات المكتوبة بخط داكن معنوية عند المستوى 0.05، وباقي المعاملات ليست معنوية الفئة المحذوفة في النموذج هي الأبيض، الذكور، SAT< 1000، 90% الأخيرة في فئة المدارس الثانوية SES المتوسطة، SEL-3، المعاهد.

معدلات التخرج هي 6 سنوات، معدلات التخرج من المدرسة الأولى وكلاهما معرفة في ملاحظات الملحق جدول (1.3.D).

الطبقات المعهدية المختارة معرفة في ملاحظات ملحق جدول (1.3.D). انظر ملحق B لتعريفات الحالة الاجتماعية الاقتصادية (SES).

SEL-1= معاهد بمتوسط درجة SAT المختلطة مساوية 1300 فأكثر.

SEL-2= معاهد بمتوسط درجة SAT المختلطة بين 1150 و 1299.

SEL-3= معاهد بمتوسط درجة SAT المختلطة بين 1150 .

المصدر: Bowen and Bok, op, cit. p.381

(a) ما هي النتائج العامة التي يمكن استنتاجها عن معدلات التخرج من كل الطلبة المسجلين بالجامعة، ومن الطلبة المسجلين بالجامعة؟

(b) نسبة الأوزان هي النسبة بين وزنين. قارن بين مجموعتين من كل المسجلين بالجامعة. مجموعة يكون فيها درجة SAT أكثر من 1299، والمجموعة الأخرى تكون درجة SAT أقل من 1000 (الفئة الأساسية). نسبة الأوزان المساوية لـ 1.393 تعني أن أوزان المسجلين بالجامعة ويتخرجون في الجامعة ويكونون من المجموعة الأولى هي 39 في المائة أعلى من نظيرهم من المجموعة الثانية. هل من الممكن تفسير نسب الأوزان المختلفة والموجودة في الجدول؟

(c) ماذا عن المعنوية الإحصائية للمعلمات المقدرة؟ وماذا عن المعنوية الكلية للنموذج المقدر؟

12.15 في نموذج البروبيت المعطى في جدول (11.15) مقدار الخطأ u_i له التباين التالي:

$$\sigma_u^2 = \frac{P_i(1 - P_i)}{N_i f_i^2}$$

بحيث إن f_i هي دالة احتمال التوزيع الطبيعي القياسي محسوبة عند $F\text{-}1(P_i)$.

(a) باستخدام تباين u_i السابقة، كيف يمكن تحويل النموذج الموجود في جدول (10.15) حتى نحصل على مقدار خطأ ثابت التباين؟

(b) استخدم بيانات جدول (10.15) لتوضيح البيانات المحولة.

(c) قدر نموذج البروبيت باستخدام البيانات المحولة، وقارن النتائج مع تلك التي تم الحصول عليها باستخدام البيانات الأصلية.

13.15 بما أن R^2 كمقياس لجودة التوفيق لاتعتبر مناسبة بشكل جيد للنماذج ذات المتغير التابع الثنائي. فقد تم اقتراح استخدام اختبار X^2 التالي:

$$\chi^2 = \sum_{i=1}^{G} \frac{N_i(\hat{P}_i - P_i^*)^2}{P_i^*(1 - P_i^*)}$$

بحيث إن:

N_i = عدد المشاهدات في الخانة i.

$\hat{P}_i$ = الاحتمال الفعلي لوقوع الحدث $(=n_i/N_i)$

P_I^* = الاحتمال المقدر

G = عدد خانات الجدول (بمعنى عدد المستويات التي يقاس فيها المتغير X، مثلاً 10 في جدول (4.15))

من الممكن ملاحظة أنه للعينات كبيرة الحجم، فإن X^2 يكون له توزيع كاي X^2 بدرجات حرية $(G-k)$ بحيث إن k هي عدد المعلمات في النموذج المقدر $(k < G)$ استخدم اختبار X^2 السابق على انحدار (1.7.15) وعلق على النتائج الخاصة بجودة التوفيق وقارنها مع قيمة R^2 المذكورة.

14.15 جدول (20.15) يعطي البيانات الخاصة بنتائج رش rotenone تركيزات مختلفة على حشرة الأقحوان (زهرة الذهب) في حوالي 50 مجموعة. أوجد نموذجاً مناسبًا للتعبير عن احتمال الوفاة كدالة في لوغاريتم X، لوغاريتم الجرعة علق على النتائج، احسب أيضاً اختيار X^2 لجودة التوفيق الذي تمت مناقشته في تمرين 13.15.

15.15 40 متقدماً لبرنامج دراسات عليا لهم درجات في الجزءين العددي واللغوي في امتحان GRE، هذه الدرجات معطاة في جدول (12.15). ستة طلاب تقدموا لبرنامج الدراسات العليا.

جدول (20.15) دراسة التسمم و rotenone على حشرة الأقحوان

Concentration, milligrams per liter				
X	$\log(X)$	Total, N_i	Death, n_i	$\hat{P}_i = n_i / N_i$
2.6	0.4150	50	6	0.120
3.8	0.5797	48	16	0.333
5.1	0.7076	46	24	0.522
7.7	0.8865	49	42	0.857
10.2	1.0086	50	44	0.880

المصدر: D.J. Fennet, probit analysis Cambridge university press, London, 1964.

جدول (21.15)

	GRE aptitude test scores		
Student number	Quantitative, Q	Verbal, V	Admitted to graduate program (Yes = 1, No = 0)
1	760	550	1
2	600	350	0
3	720	320	0
4	710	630	1
5	530	430	0
6	650	570	0
7	800	500	1
8	650	680	1
9	520	660	0
10	800	250	0
11	670	480	0
12	670	520	1
13	780	710	1

المصدر: Donald F.Marrison, Applied linear statical method, prentice- hall, inc., Englewood cliffs, N.J., 1983 p.279.

(a) استخدم نموذج LPM للتنبؤ باحتمال التقدم لبرنامج الدراسات العليا معتمدين على درجات الجزء العددي واللغوي في امتحان الـ GRE.

(b) هل هذا النموذج مناسب وكاف؟ إذاكانت الإجابة بلا، ما هي البدائل التي يمكن أن تقترحها؟

16.15 لدراسة فعالية كوبونات تخفيض السعر على عبوة من ست زجاجات من مشروب غازي (2 لتر)، Douglas Montgomery, Elizabeth Peck دوجلاس مونتجومري وإليزابيت بيك قاموا بتجميع بيانات موضحة في جدول (22.15). العينة من 5500 مستهلك تم اختيارها عشوائياً وتوزيعها على أحد عشر نوعاً من الأنواع المختلفة لهذه الكوبونات، ولكل نوع يوجد 500 مشاهدة موضحة في الجدول. المتغير التابع في هذه الدراسة هو ما إذاكان المستهلك سيستخدم الكوبون في خلال شهر من الشراء أولا.

(a) هل نموذج اللوجيت مناسب لهذه البيانات؟ استخدم معدل استخدام الكوبون كالمتغير التابع، ومقدار التخفيض في السعر كالمتغير المفسر.

(b) ادرس ما إذا كان نموذج البروبيت له نفس القدرة على تفسير البيانات كنموذج اللوجيت أم لا.

(c) ما هي القيمة المتوقعة لمعدل استخدام الكوبون إذا كان التخفيض في السعر 17 سنتاً.

(d) قدر التخفيض في السعر إذا علمت أن نسبة 70 % من الكوبونات سيتم استخدامها.

جدول (22.15)

Price discount X, ¢	Sample size N_i	Number of coupons redeemed n_i
5	500	100
7	500	122
9	500	147
11	500	176
13	500	211
15	500	244
17	500	277
19	500	310
21	500	343
23	500	372
25	500	391

المصـــدر: Douglas C.Montgomeru, Elizabeth A.Peck, introduction to linear regression analysis, John Wiley& Sons, New York, 1982, p243. (الترميز مختلف)

17.15 لمعرفة ما إذا كان الفرد لديه حساب في البنك (جاري، توفير، . .) من عدمه. جون كاسكي وأندروا بيرسون قدرا نموذج بروبيت للفترة 1977 إلى 1989 مستخدمين بيانات عن الأسر في الولايات المتحدة الأمريكية. النتائج معطاة في جدول (23.15). قيم معاملات الانحدار المعطاة في الجدول تقيس التأثير الحادث لكل وحدة تغير في المتغير المنحدر على احتمال أن تكون الأسرة لها حساب في البنك، هذه التأثيرات الحدية محسوبة عند القيم المتوسطة للمتغيرات المنحدرة الموجودة في النموذج.

(a) في عام 1977؟ ما تأثير الحالة الاجتماعية على وجود حساب في البنك للأسرة؟ وكذلك في عام 1989؟ هل هذه النتائج ممكن فهمها اقتصادياً؟

(b) لماذا يأخذ معامل متغير الأقلية قيمًا سالبة في كل من عامي 1977 و 1989؟

(c) كيف يمكنك تفسير الإشارة السالبة لمتغير عدد الأطفال؟

(d) ماذا تقترح قيمة إحصاء مربع كاي X^2 المعطاة في الجدول؟

(ملاحظة: تمرين 13.15)

18.15 دراسة Monte Carlo: كوسيلة لفهم نموذج البروبيت، افترض وليم بيكر ودونالد ودمان التالي(*):

$$E(Y \mid X) = -1 + 3X$$

بوضع $Y_i = t + 3X + \varepsilon_i$ بحيث إن ε_i مفترض أنه يتبع التوزيع الطبيعي القياسي (بمعنى أن: التوقع = صفر و التباين = 1)، قاما بتخليق عينة من 35 مشاهدة، كما هو موضح في جدول (24.15).

(a) من بيانات Y و X المعطاة في الجدول. هل من الممكن أن نقدر LPM؟ تذكر أن القيم الحقيقية لـ $E(Y|X) = -1 + 3X$.

(b) بمعلومية $X = 0.48$ قدر $E = (Y \mid X = 0.48)$ وقارن ذلك مع القيمة الحقيقية لـ $E(Y \mid X = 0.48)$. لاحظ $\bar{X} = 0.48$.

(*) William E.Becker and Donald M. Waldman, "A graphical interpretation of probit coefficients", journal of economic education, vol. 20, no.4, fall 1989, pp. 371- 378.

جدول (23.15) الأرقام بين الأقواس هي قيم إحصاء t .

	1977 data		1989 data	
	Coefficients	Implied slope	Coefficients	Implied slope
Constant	−1.06 (3.3)*		−2.20 (6.8)*	
Income (thousands 1991 $)	0.030 (6.9)	0.002	0.025 (6.8)	0.002
Married	0.127 (0.8)	0.008	0.235 (1.7)	0.023
Number of children	−0.131 (3.6)	−0.009	−0.084 (2.0)	−0.008
Age of head of household (HH)	0.006 (1.7)	0.0004	0.021 (6.3)	0.002
Education of HH	0.121 (7.4)	0.008	0.128 (7.7)	0.012
Male HH	−0.078 (0.5)	−0.005	−0.144 (0.9)	−0.011
Minority	−0.750 (6.8)	−0.050	−0.600 (6.5)	−0.058
Employed	0.186 (1.6)	0.012	0.402 (3.6)	0.039
Homeowner	0.520 (4.7)	0.035	0.522 (5.3)	0.051
Log likelihood	−430.7		−526.0	
Chi-square statistic (H_0: All coefficients except constant equal zero)	408		602	
Number of observations	2,025		2,091	
Percentage in sample with correct predictions	91		90	

المصـدر: John P.Caskey and Andrew Peterson, "Who has a bank account and who who doesn't: 1977 and 1989" Research working paper 93-10, Federal Reserve bank of Kanses city, October 1993.

(c) باستـخـدام البـيـانات عن Y^* و X المعطاة في جـدول (24.15) قـدر نموذج البروبيت، استخدم حزمة التحليل الإحصائي التي ترغب فيها. الكاتب قدر نموذج البروبيت كالتالي:

$$\hat{Y}_i^* = -0.969 + 2.764X_i$$

أوجد قيمة $P(Y^* = 1 \mid X = 0.48)$ وهذا هو $P(Y_1 > 0 \mid X = 0.48)$ هل إجابتك تتفق مع إجابة الكاتب وهي 0.64؟

(d) الانحراف المعياري للعينة لقيم المتغير X المعطاة في جدول (24.15) هو 0.31. ما هو تقدير التغير في الاحتمال إذا كانت الـ X أعلى من القيمة المتوسطة لواحد انحراف معياري، أو بعبارة أخرى، ما هي قيمة $P(Y^* = 1\ (0.79)$؟ إجابة الكاتب هي 0.25.

جدول (24.15) بيانات افتراضية مخلقة للنموذج $Y = -1 + 3X + \varepsilon$ و $Y^* = 1$ إذا كانت $Y > 0$

Y	Y*	X	Y	Y*	X
−0.3786	0	0.29	−0.3753	0	0.56
1.1974	1	0.59	1.9701	1	0.61
−0.4648	0	0.14	−0.4054	0	0.17
1.1400	1	0.81	2.4416	1	0.89
0.3188	1	0.35	0.8150	1	0.65
2.2013	1	1.00	−0.1223	0	0.23
2.4473	1	0.80	0.1428	1	0.26
0.1153	1	0.40	−0.6681	0	0.64
0.4110	1	0.07	1.8286	1	0.67
2.6950	1	0.87	−0.6459	0	0.26
2.2009	1	0.98	2.9784	1	0.63
0.6389	1	0.28	−2.3326	0	0.09
4.3192	1	0.99	0.8056	1	0.54
−1.9906	0	0.04	−0.8983	0	0.74
−0.9021	0	0.37	−0.2355	0	0.17
0.9433	1	0.94	1.1429	1	0.57
−3.2235	0	0.04	−0.2965	0	0.18
0.1690	1	0.07			

المصـــــدر: Willian E.Becker and Donald M.Waldman, "Agraphical Interpretation of probit coefficients". Journal of economic education, fall 1989, table, p.373.

APPENDIX ملحق A -15

1. A 15 تقديرات الإمكان الأعظم لنماذج البروبيت واللوجيت في حالة البيانات المفردة (غير المجمعة)(*)

MAXIMUM LIKELIHOOD ESTIMATION OF THE LOGIT AND PROBIT MODELS FOR INDIVIDUAL (UNGROUPED) DATA

كما سبق، دعنا نفترض أننا نريد تقدير احتمال أن تمتلك الأسرة المنزل السكني وفقاً لمستوى الدخل X. نحن نفترض أن هذا الاحتمال من الممكن التعبير عنه في صورة دالة اللوجيتيك (2.5.15) والتي يمكن كتابتها كالتالي:

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} \qquad (1)$$

نحن في الواقع لانشاهد P_i ولكن نشاهد النتيجة $Y = 1$ إذا امتلك المفردة المنزل السكني و $Y = 0$ عندما لاتمتلك المفردة المنزل.

(*) هذه المناقشة معتمدة بشكل رئيسي على:
John Neter, Michael H. Kutner, Christopher J. Nachstein and William Wasserman, applied linear statistical models, 4th ed., Irwin, 1996, pp. 573- 574.

وبما أن Y_i هو متغير برنولي، فمن الممكن كتابة التالي:

$$\Pr(Y_i = 1) = P_i \tag{2}$$

$$\Pr(Y_i = 0) = (1 - P_i) \tag{3}$$

افترض أن لدينا عينة عشوائية من n مشاهدة، دع $f_i(Y_i)$ ترمز إلى احتمال $Y=1$ أو 0، الاحتمال المشترك لمشاهدة n قيمة من Y، بمعنى آخر $f(Y_1, Y_2, ..., Y_n)$ هو كالتالي:

$$f(Y_1, Y_2, \ldots, Y_n) = \prod_1^n f_i(Y_i) = \prod_1^n P_i^{Y_i}(1 - P_i)^{1-Y_i} \tag{4}$$

بحيث إن $\prod$ هي علامة حاصل الضرب. لاحظ أننا نستطيع كتابة دالة الاحتمال المشتركة كحاصل ضرب دوال الاحتمال الفردية، حيث إن Y_i مشاهدات مستقلة وكل Y_i لها نفس دالة الاحتمال لتوزيع اللوجيستيك. دالة الاحتمال المشتركة الموجودة في المعادلة (4) معروفة باسم دالة الإمكان (LF).

المعادلة (4) من الصعب التعامل معها، ولكن إذا أدخلنا عليها اللوغاريتم الطبيعي، نحصل على ما يعرف باسم لوغاريتم دالة الإمكان (LLF)

$$\begin{aligned}\ln f(Y_1, Y_2, \ldots, Y_n) &= \sum_1^n [Y_i \ln P_i + (1 - Y_i) \ln (1 - P_i)] \\ &= \sum_1^n [Y_i \ln P_i - Y_i \ln (1 - P_i) + \ln (1 - P_i)] \\ &= \sum_1^n \left[Y_i \ln \left(\frac{P_i}{1 - P_i}\right)\right] + \sum_1^n \ln (1 - P_i)\end{aligned} \tag{5}$$

من (1) من السهل اثبات أن:

$$(1 - P_i) = \frac{1}{1 + e^{\beta_1 + \beta_2 X_i}} \tag{6}$$

وأيضاً:

$$\ln \left(\frac{P_i}{1 - P_i}\right) = \beta_1 + \beta_2 X_i \tag{7}$$

باستخدام (6) و (7) من الممكن كتابة الـLLF الموجودة في معادلة (5) كالتالي:

$$\ln f(Y_1, Y_2, \ldots, Y_n) = \sum_1^n Y_i(\beta_1 + \beta_2 X_i) - \sum_1^n \ln \left[1 + e^{(\beta_1 + \beta_2 X_i)}\right] \tag{8}$$

كما ترى من المعادلة (8) لوغاريتم دالة الإمكان هو دالة في المعاملات β_1 و β_2، بما أن قيم X_i هي قيم معلومة.

في ML هدفنا هو تعظيم الـ LF (أو LLF) وذلك للحصول على قيم المعلمات المجهولة بحيث يكون احتمال مشاهدة قيم الـ Y's كبير (معظم) بقدر الإمكان. ولهذا الغرض نحصل على التفاضل الجزئي للمعادلة (8) بالنسبة للمعلمات المجهولة. ثم من الممكن تطبيق الشرط الثاني (المشتقة التفاضلية الثانية) للتعظيم لإثبات أن قيم المعلمات التي تم الحصول عليها هي في الحقيقة القيم العظمى لـ LF. وبالتالي نحتاج لتفاضل (8) بالنسبة لـ β_1 و β_2. وستدرك سريعاً أن النتيجة ستكون غير خطية بشكل واضح في المعالم، ولايوجد حل واضح من الممكن الحصول عليه. ولهذا السبب سنحتاج لاستخدام إحدى طرق التقدير غير الخطية التي تعرضنا لها في الفصل السابق للحصول على حلول عددية. وبمجرد الحصول على قيم عددية لـ β_1 و β_2 من الممكن بسهولة تقدير (1).

طريقة ML لنموذج البروبيت مشابهة لنظيرها في نموذج اللوجيت باستثناء أن (1) تستخدم فيها دالة التوزيع التراكمية للتوزيع الطبيعي وليس CDF للتوزيع اللوجيستيكي. النتيجة الحاصل عليها ستكون أكثر تعقيداً، ولكن الفكرة العامة واحدة، ولذلك لن نتطرق إليها أكثر من ذلك.

الفصل السادس عشر

نمـــاذج انحـــدار البيـــانات الطولية

PANEL DATA REGRESSION MODELS

في الفصل الأول، ناقشنا أنواع البيـانات المستـخدمة بوجـه عام في التـحليل التطبيقي مثل السلاسل الزمنية، البيانات المقطعية والبيانات الطولية. في بيانات السلاسل الزمنية نشاهد قيم متغير واحد أو أكثر في خلال فترة زمنية ما. (مثال GPA لفترات ربع سنوية أو سنوية). في البيانات المقطعية بيانات متغير واحد أو أكثر من متغير يتم تجميعها للعديد من وحدات العينة (مثال معدل الجريمة في الولايات الخمسين في الولايات المتحدة الأمريكية لسنة ما) في البيانات الطولية تكون وحدة البيانات عينة قطاع عرضي (مثلاً أسرة، مؤسسة، أو ولاية) يتم مسحها خلال فترة زمنية ما. باختصار البيانات الطولية لها بعد زمن وبعد مكاني أيضاً.

قد رأينا بالفعل مثالاً على ذلك في جدول (1.1)، والذي يحتوي على بيانات عن إنتاج البيض وأسعاره في 50 ولاية في الولايات المتحدة في عامي 1990 و 1991. لأي سنة معينة فإن البيانات الخاصة بالبيض وأسعاره تمثل عينة بيانات مقطعية. وبالنسبة لولاية معينة، فإن هناك مشاهدتين لسلاسل زمنية واحدة خاصة بإنتاج البيض، والأخرى خاصة بالسعر. وبالتالي مجموع المشاهدات (المجمعة) لدينا هي (50 x 2) = 100 مشاهدة على إنتاج البيض وأسعاره.

هذه المشاهدات تسمى بيانات طولية أي بيانات تجميعية (تجميع من السلاسل الزمنية ومشاهدات مقطعية) أو خليط من بيانات السلاسل الزمنية والبيانات المقطعية (دراسة لمتغير أو مجموعة من المواضيع خلال فترة زمنية) أو تحليل الأحداث التاريخية (مثال دراسة تحركات سياسية ما خلال فترة زمنية خاصة) وتسمى أيضًا التحليل الجماعي (مثال تتبع التطور الوظيفي لـ 1965 متخرج في كليات التجارة). وعلى الرغم من الاختلافات البسيطة بين هذه المسميات، فإنها جميعًا في النهاية تدرس

بيانات مقطعية وتحركاتها خلال فترة زمنية ما. وسنقوم باستخدم مصطلح البيانات الطولية للتعبير عن هذه الحالة من البيانات.

وسنطلق على نموذج الانحدار الذي يعتمد على مثل هذه البيانات الطولية باسم نماذج انحدار البيانات الطولية .

البيانات الطولية يتم استخدامها بشكل متزايد في الأبحاث الاقتصادية، بعض هذه أنواع البيانات الطولية المعروفة هي:

1 - الدراسة الطولية لتحركات الدخل (PSID) والتي قام بها مركز الأبحاث الاجتماعية في جامعة ميتشيجن. بدأت الدراسة في عام 1968، حيث يقوم المركز في كل عام بتجميع بيانات عن 5000 أسرة، هذه البيانات خاصة بمتغيرات اقتصادية وسكانية متعددة.

2 - مكتب التعداد السكني التابع للقسم التجاري، قام بمسح مماثل لـ PSID يسمى مسح الدخل والمشاركة في البرنامج (SIPP). يقوم المشتركون في هذه الدراسة بالإجابة عن أسئلة خاصة بأوضاعهم الاقتصادية أربع مرات كل عام.

هناك العديد من المسوح الأخرى التي تقوم بها وكالات حكومية متعددة. ولابد من التنبيه على شئ مهم هنا. فموضوع انحدار البيانات الطولية هو موضوع شديد الاتساع والخلفية الرياضية والإحصائية الخاصة به تعتبر معقدة نوعًا ما، وبالتالي نحن نتمنى أن نتطرق إلى أساسيات الموضوع تاركين التفاصيل لكتب ومراجع أخرى[1].

مع مراعاة أن بعض هذه المراجع له مستوى فني عال. لحسن الحظ، فإن بعض حزم البرامج الإلكترونية مثل Eviews و Pc Give, SAS, STATA, Shazam و Limdep وآخرين جعلوا مهمة انحدار البيانات الطولية سهلة نسبيًا.

(1) بعض المراجع هي: G. Chamberlain, "Panel Data", in Handbook of Econometrics, vol. II, Z. Griliches and M. D. Intriligator, ed., North-Holland Publishers, 1984, Chap. 22.; C. Hsiao, Analysis of Panel Data, Cambridge University Press, 1986; G. G. Judge, R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. C. Lee, Introduction to the Theory and Practice of Econometrics, 2d ed., John Wiley & Sons, New York, 1985, Chap. 11; W. H. Greene, Econometric Analysis, 4th ed., Prentice-Hall, Englewood Cliffs, N. J., 2000, Chap. 14; Badi H. Baltagi, Econometric Analysis of Panel Data, John Wiley and Sons, New York,1995' and J. M. Wooldridge, Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, Mass., 1999.

1.16 لماذا تستخدم البيانات طولية؟ WHY PANEL DATA?

ما هي مميزات البيانات الطولية، والتي تميزها عن بيانات السلاسل الزمنية أو البيانات المقطعية؟ Baltagi استعرض المميزات التالية للبيانات الطولية (2).

1 - بما أن البيانات طولية مرتبطة بالمفردات، والمؤسسات، الولايات، البلدان وهكذا خلال فترة زمنية، فمن المحتمل وجود عدم تجانس بين هذه الوحدات. الأسلوب المستخدم للتقدير في البيانات يأخذ مسألة عدم التجانس في الاعتبار، بحيث يسمح بوجود متغيرات مجددة للمفردات، كما سنرى لاحقًا. سنستخدم المصطلح مفردة للتعبير عن الوحدة الجزئية مثل المفردات، المؤسسات، الولايات، والبلدان المختلفة.

2 - عندما يتم المزج بين بيانات السلاسل الزمنية، والبيانات المقطعية، نحصل على البيانات الطولية، وبالتالي فهي تعطي معلومات أكثر عن البيانات بتباين أكثر وأقل إرتباط تداخلي بين المتغيرات ودرجات حرية أكثر وكفاءة أكثر.

3 - بدراسة البيانات المتكررة المقطعية تكون البيانات الطولية مناسبة أكثر لدراسة حركية التغير. معدلات البطالة، والتحول الوظيفي، وتحرك قوة العمل من الأفضل دراستها من خلال البيانات الطولية.

4 - البيانات الطولية من الممكن أن تتنبأ وتقيس التأثيرات التي لاتستطيع ببساطة مشاهدتها من خلال البيانات المقطعية فقط، أو بيانات السلاسل الزمنية فقط. من الأفضل استخدامها عند دراسة التحركات الناجحة لزيادة الحد الأدنى للأجر من خلال الحد الأدنى للأجر بالولاية أو الاتحاد.

5 - البيانات الطولية تجعل من الممكن دراسة النماذج السلوكية الأكثر تعقيدًا. فمثلاً ظاهرة مثل اقتصاديات القياس، والتغيرات التكنولوچية، من الأفضل دراستها من خلال البيانات الطولية عن دراستها من خلال بيانات مقطعية فقط، أو بيانات سلاسل زمنية فقط.

6 - عندما تكون البيانات متاحة للعديد من آلاف الوحدات، يمكن أن تقلل بيانات الطولية من التحيز الذي قد يتواجد في النتائج إذا قمنا بتجميع المفردات أو المؤسسات في تجميعة واحدة.

باختصار، فإن البيانات الطولية تزيد من جودة التحليل الاختياري بطريقة قد لا تكون ممكنة إذا استخدمنا بيانات مقطعية فقط، أو بيانات السلاسل الزمنية فقط. وهذا لا يعني أنه لاتوجد مشاكل في استخدام البيانات الطولية، وسنقوم باستعراض هذه المشاكل من الجانب النظري، بعد أن نناقش مثالاً عن البيانات الطولية.

---

(2) Battagi, op.cit., pp.3-6.

2.16 البيانات الطولية.. مثال توضيحي :
PANEL DATA: AN ILLUSTRATIVE EXAMPLE

لتوضيح الموضوع، دعنا نستعرض مثالاً واقعياً. البيانات الموجودة في جدول (1.16) مأخوذة من دراسة شهيرة في نظرية الاستثمار مقدمة من جرنفيلد[3].

جرنفيلد اهتم بمعرفة كيف يعتمد النمو الاستثماري الحقيقي (Y) على القيمة الحقيقية للمؤسسة X_2 وأسهم رأس المال الحقيقية X_3.

وعلى الرغم من أن الدراسة شملت العديد من الشركات، فإنه لسهولة التوضيح حصلنا على بيانات عن أربع شركات فقط هي جنرال إلكتريك (GE)، جنرال موتورز (GM)، الولايات المتحدة للصلب (US) وويستنجهوس.

البيانات المتاحة عن كل شركة خاصة بالمتغيرات الثلاثة السابق ذكرها ومتاحة خلال الفترة 1935–1945، وبالتالي فإن لدينا أربع وحدات مقطعية و 20 فترة زمنية. الإجمالي هو 80 مفردة. مسبقًا فإن من المتوقع أن تكون y على علاقة طردية (موجبة) مع كل من X_2، X_3.

في الأصل من الممكن أن نقوم بأربعة انحدارات سلاسل زمنية، واحد لكل شركة أو من الممكن أن نقوم بعشرين انحداراً مقطعي، واحد لكل سنة، مع الوضع في الاعتبار أنه في الحالة الأخيرة ستكون هناك مشكلة خاصة بدرجات الحرية[4].

---

(3) Y. Grunfeld, "The Determinants of Corporate Investment," unpublished Ph.D. thesis, Department of Economics, University of Chicago, 1958. The data are reproduced in several books. We have taken them from H. D. Vinod and Aman Ullha, Recent Advances in Regression Methods, Marcel Dekker, New York, 1981, pp. 259-261. The Grunfeld study has become a favorite of texttbook writers as the data is manageable for illustraion purposes.

(4) لكل سنة لدينا أربعة مشاهدات فقط عن المتغير المنحدر عليه والمتغيرات المنحدرة. وإذا أدخلنا الجزء المقطوع من المحور الصادي في الاعتبار سيكون لدينا ثلاثة معالم للتقدير بدرجة حرية واحدة. بالطبع مثل هذا الانحدار قد يكون بدون معنى.

جدول (1.16) بيانات الاستثمار في أربع شركات ، 1935– 1954

Observation	I	F_{-1}	C_{-1}	Observation	I	F_{-1}	C_{-1}
GE				US			
1935	33.1	1170.6	97.8	1935	209.9	1362.4	53.8
1936	45.0	2015.8	104.4	1936	355.3	1807.1	50.5
1937	77.2	2803.3	118.0	1937	469.9	2673.3	118.1
1938	44.6	2039.7	156.2	1938	262.3	1801.9	260.2
1939	48.1	2256.2	172.6	1939	230.4	1957.3	312.7
1940	74.4	2132.2	186.6	1940	361.6	2202.9	254.2
1941	113.0	1834.1	220.9	1941	472.8	2380.5	261.4
1942	91.9	1588.0	287.8	1942	445.6	2168.6	298.7
1943	61.3	1749.4	319.9	1943	361.6	1985.1	301.8
1944	56.8	1687.2	321.3	1944	288.2	1813.9	279.1
1945	93.6	2007.7	319.6	1945	258.7	1850.2	213.8
1946	159.9	2208.3	346.0	1946	420.3	2067.7	232.6
1947	147.2	1656.7	456.4	1947	420.5	1796.7	264.8
1948	146.3	1604.4	543.4	1948	494.5	1625.8	306.9
1949	98.3	1431.8	618.3	1949	405.1	1667.0	351.1
1950	93.5	1610.5	647.4	1950	418.8	1677.4	357.8
1951	135.2	1819.4	671.3	1951	588.2	2289.5	341.1
1952	157.3	2079.7	726.1	1952	645.2	2159.4	444.2
1953	179.5	2371.6	800.3	1953	641.0	2031.3	623.6
1954	189.6	2759.9	888.9	1954	459.3	2115.5	669.7
GM				WEST			
1935	317.6	3078.5	2.8	1935	12.93	191.5	1.8
1936	391.8	4661.7	52.6	1936	25.90	516.0	0.8
1937	410.6	5387.1	156.9	1937	35.05	729.0	7.4
1938	257.7	2792.2	209.2	1938	22.89	560.4	18.1
1939	330.8	4313.2	203.4	1939	18.84	519.9	23.5
1940	461.2	4643.9	207.2	1940	28.57	628.5	26.5
1941	512.0	4551.2	255.2	1941	48.51	537.1	36.2
1942	448.0	3244.1	303.7	1942	43.34	561.2	60.8
1943	499.6	4053.7	264.1	1943	37.02	617.2	84.4
1944	547.5	4379.3	201.6	1944	37.81	626.7	91.2
1945	561.2	4840.9	265.0	1945	39.27	737.2	92.4
1946	688.1	4900.0	402.2	1946	53.46	760.5	86.0
1947	568.9	3526.5	761.5	1947	55.56	581.4	111.1
1948	529.2	3245.7	922.4	1948	49.56	662.3	130.6
1949	555.1	3700.2	1020.1	1949	32.04	583.8	141.8
1950	642.9	3755.6	1099.0	1950	32.24	635.2	136.7
1951	755.9	4833.0	1207.7	1951	54.38	732.8	129.7
1952	891.2	4924.9	1430.5	1952	71.78	864.1	145.5
1953	1304.4	6241.7	1777.3	1953	90.08	1193.5	174.8
1954	1486.7	5593.6	2226.3	1954	68.60	1188.9	213.5

ملاحظات :

$I = Y$ =النمو الاستثماري= بالإضافة إلى الأرض والمعدات والصيانة والتصليح بملايين الدولارات منخفضة بـ P_1.

$F = X_2$ =قيمة المؤسسة= سعر الحصص العادية والمفضلة في 31 ديسمبر (أو متوسط السعر لـ 31 ديسمبر و31 يناير للسنة التالية) مضروبًا في عدد الحصص العادية والمفضلة الباقية مضافًا إليه القيمة الكلية الكتابية للأصل في 31 ديسمبر بملايين الدولارات منخفضة بـ P_2.

$C = X_3$ = أسهم الأرض والمعدات= المجموع التراكمي للزيادة الصافية للأرض والمعدات منخفضة بـ P_1 مطروحة من. . .

منخفضة بـ P_3 في هذه التعريفات.

P_1 = مخفض ضمني للسعر الخاص بالمعدات الإنتاجية المتينة (100= 1947)

P_2 = مخفض ضمني بسعر GNP (100= 1947)

P_3 = مخفض تكلفة= مؤشر لمتوسط سعر البيع بالجملة لمدة 10 سنوات للمعدن والمنتجات المعدنية (1947=100).

المصـــدر : Reproduction from H.D. Vinod and Aman Vallen, Recent advances in regression methods, Marcel Dekker, New York, 1981, pp. 259-261.

مزج أو تجميع الثمانين مفردة، يمكننا من كتابة دالة جرينفيلد Grunfeld للاستثمار كالتالي:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it}$$
$$i = 1, 2, 3, 4 \qquad (1.2.16)$$
$$t = 1, 2, \dots, 20$$

بحيث إن i ترمز إلى الوحدة رقم i في البيان المقطعي، و t ترمز إلى الفترة الزمنية رقم t. للتسهيل سنجعل i ترمز إلى البيان المقطعي، و t ترمز للزمن. من المفترض أن أكبر عدد لوحدات أو مشاهدات المقطعية هو N، وأكبر عدد من الفترات الزمنية هو T. إذا كانت كل وحدة مقطعية لها نفس العدد من مشاهدات السلاسل الزمنية تسمى البيانات الطولية بالبيانات الطولية المتوازنة. في المثال الحالي لدينا بيانات طولية متوازنة، حيث إن كل شركة في العينة لديها 20 مشاهدة (مفردة). إذا كان عدد المشاهدات يختلف داخل الوحدات، فإننا نسمي تلك البيانات بالبيانات الطولية غير المتوازنة.

في هذا الفصل، سينصب معظم اهتمامنا على البيانات الطولية المتوازنة، مبدئيًا دعنا نفترض أن X's ليست عشوائية ومقدار الخطأ يتبع الفروض التقليدية، بمعنى أن $E(u_{it})\sim N(0, \sigma^2)$. لاحظ أن الترميزين الثنائي والثلاثي يعتمدان على طبيعة المسألة المراد تفسيرها.

كيف سنقوم بتقدير (1.2.16)؟ الإجابة كالتالي:

3.16 تقدير نماذج انحدار البيانات الطولية.. أسلوب التأثيرات الثابتة: ESTIMATION OF PANEL DATA REGRESSION MODELS THE FIXED EFFECTS APPROACH

تقدير المعادلة (1.2.16) يعتمد على الفروض المتعلقة بالجزء المقطوع من المحور الصادي، ومعاملات الانحدار (الميل) ومقدار الخطأ u_{it}، هناك العديد من الحالات الممكنة لهذه الفروض المختلفة[5]:

1 – افترض أن الجزء المقطوع من المحور الصادي، ومعاملات الانحدار هي ثوابت بالنسبة للزمن والمشاهدات. ومقدار الخطأ يحتوي على الفروق الناتجة عن اختلاف الزمن والمشاهدات.

(5) هذا الشرح معتمد على Judge etal., op.cit., and Hsiao, op.cit., pp. 9-10.

2 – معـامـلات الانحـدار ثوابت، ولكن الجـزء المقطوع من المحـور الصـادي يخـتلف باختلاف المشاهدات.

3 – معـامـلات الانحـدار ثوابت، ولكن الجـزء المقطوع من المحـور الصـادي يخـتلف باختلاف المشاهدات والزمن.

4 – كل المعـامـلات (الجـزء المقطوع مـن المحـور الصـادي ومـعـامـلات الميل) تخـتلف باختلاف المشاهدات.

5 – الجزء المقطوع من المحور الصـادي ومعـاملات الانـحدار (الميل) تختلف باختـلاف المشاهدات والزمن.

كما نرى، فإن كـلاً من هذه الحالات، تعتبر تعقيدًا أكثر (وربما واقـعية أكثر) في تقدير نماذج انحدار البـيانات الطوليـة مثل المذكـورة في (1.2.16). بالطبع سيـزداد التـعقيـد بإضـافـة المزيد من المتغيـرات المنحـدرة للنمـوذج، حيـث تكـون هناك إمكانيـة للارتباط الخطي بين هذه المتغيرات.

للتـعـمق في كل حـالة من الحـالات السـابق ذكرهـا يحتـاج على كـتـاب منفـصل، وهناك بالفعـل العـديـد من الكتـب في ذلك المجال[6]. فيـمـا يلي سنقـوم باسـتـعـراض بعض الصفـات الرئيسية للحـالات السابقة، مع التركيـز على الحالات الأربع الأولى. مناقشتنا للموضوع ليس لها جانب فني.

1 - كل المعاملات ثابتة بالنسبة للزمن والمشاهدات:

1 –All coefficients constant across time and individuals

الأبسط وربما يكـون افـتـراضًـا سـاذجًـا غيـر واقـعي، هو عـدم الاعـتـبار لاخـتـلاف المشاهدات، واختـلاف الزمن للبيانات المجـمعة، والتقـدير الانحداري عن طريق OLS العادية. وذلك يتم من خـلال وضع الـ 20 مشـاهـدة الخاصـة بكل شـركـة فـوق بعضـهـا البـعض، وبالتـالي نحـصل على 80 مـشاهـدة لكـل متـغيـر من المتـغيـرات المـوجـودة في النموذج. نتائج الـ OLS لمثالنا الحالي جاءت كالتالي:

(6) بالإضافة إلى الكتب المذكورة في الملاحظة 1. انظر أيضًا:

Terry F.Dielman, Pooled cross- sectional and time series data analysis, Marcel Dekker, New york, 1989, and Lois W.Sayrs, pooled time series analysis, sage publications, New bury park, California, 1989.

$$\begin{aligned} \hat{Y} &= -63.3041 + 0.1101X_2 + 0.3034X_3 \\ se &= (29.6124) \quad (0.0137) \quad (0.0493) \\ t &= (-2.1376) \quad (8.0188) \quad (6.1545) \\ R^2 &= 0.7565 \quad \text{Durbin–Watson} = 0.2187 \\ n &= 80 \quad df = 77 \end{aligned} \qquad (1.3.16)$$

إذا استخدمنا الطريقة العادية للتعليق على نتائج هذا الانحدار التجميعي، نجد أن كل معاملات الانحدار لها معنوية إحصائية منفردة. وكل معاملات الانحدار لها الإشارة الموجبة المتوقعة، وقيمة R^2 هي قيمة مرتفعة إلى حد كبير. كما هو متوقع، فإن Y على علاقة موجبة مع X_2 و X_3. النتيجة الوحيدة المنخفضة نوعًا ما هي قيمة إحصاء داربن وتسون Durbin-watson وهذه القيمة المنخفضة قد تكون دليلاً على وجود ارتباط ذاتي في البيانات. بالطبع نحن نعلم أن قيمة درابن وتسون المنخفضة قد تكون نتيجة خطأ في التوصيف. فعلى سبيل المثال، النموذج المقدر يفترض أن الجزء المقطوع من المحور الصادي متساوي لكل من GE و GM و US و Westinghouse وأيضًا يفترض أن معاملات الميل لكل من متغيرين الـ X متساويًا تمامًا لكل من الشركات الأربع السابق ذكرها. وكما هو ملاحظ، فإن هذا الفرض شديد التعقيد. وبالتالي بغض النظر عن بساطة النموذج، فإن الانحدار المجمع (1.2.16) لايستطيع التعبير عن العلاقة الحقيقية بين Y و الـ X's بالنسبة للشركات الأربع. ما نحتاجه هو إيجاد طريق ما لأخذ الاختلاف بين الشركات الأربع في الاعتبار. كيف يمكننا عمل ذلك؟ ذلك مشروح في الفقرة التالية:

2 - معاملات الميل ثابتة، ولكن الجزء المقطوع من المحور الصادي يختلف باختلاف المشاهدات: نموذج انحدار التأثيرات الثابتة أو طريقة المربعات الصغرى باستخدام المتغير الوهمي (LSDV)

2 - Slope coefficients constant but the intercept varies across individuals: The fixed effects or least-squares dummy variable (LSDV) regression model.

اختلاف الجزء المقطوع من المحور الصادي باختلاف الشركة يجعلنا نضع في الاعتبار - عند التحليل - الاختلافات الفردية لكل شركة أو كل وحدة في البيانات المقطعية، ولكننا ما زلنا نفترض أن معاملات الميل ثابتة بالنسبة لكل الشركات. لنرى ذلك، دعنا نكتب النموذج (1.2.16) كالتالي:

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \qquad (2.3.16)$$

لاحظ أننا وضعنا الترميز i على الجزء المقطوع من المحور الصادي للتعبير عن إمكانية اختلافه بين الشركات الأربع. هذه الاختلافات قد تكون بسبب صفات خاصة بكل شركة، مثل أسلوب الإدارة، أو فلسفة الإدارة. تاريخيًا، فإن النموذج (2.3.16) معروف باسم نموذج انحدار التأثيرات الثابتة (FEM). المصطلح "التأثيرات الثابتة" يعود إلى أنه على الرغم من اختلاف الجزء المقطوع من المحور الصادي بين المشاهدات (الأربع شركات) فإنه لايختلف باختلاف الزمن، وبالتالي هو ثابت زمنيًا. لاحظ أنه إذا كتبنا الجزء المقطوع من المحور الصادي على أنه β_{1it} فإن ذلك يعني أن الجزء المقطوع من المحور الصادي لكل شركة يختلف باختلاف الزمن. ونلاحظ أن FEM المعطى في (2.3.16) يفترض أن معاملات (الميل) للمتغيرات المنحدرة لاتتغير بتغير الشركات ولاتتغير مع الزمن.

كيف نسمح فعليًا للجزء المقطوع من المحور الصادي (تأثير ثابت) للاختلاف بين الشركات؟ من السهل فعل ذلك باستخدام أسلوب المتغير الوهمي الذي سبق وذكرناه في الفصل (9)، وبالأخص أسلوب الجزء المقطوع من المحور الصادي الوهمي القابل للتفاضل، وبالتالي نكتب (2.3.16) كالتالي:

$$Y_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \quad \textbf{(3.3.16)}$$

بحيث إن : $D_{2i}=1$ وإذا كانت المفردة تنتمي إلى شركة GM، 0 بخلاف ذلك، $D_{3i}=1$ إذا كانت المفردة تنتمي إلى شركة US، 0 بخلاف ذلك، $D_{ui}=1$ إذا كانت المفردة تنتمي إلى شركة WEST، 0 بخلاف ذلك. وبما أن لدينا أربع شركات، فإننا نستخدم ثلاثة متغيرات وهمية فقط لتجنب الوقوع في مشكلة المتغير الوهمي (مثل في حالة الارتباط الخطي التام). فهنا لايوجد متغير وهمي لشركة GE. بمعنى آخر، فإن α_1 تمثل الجزء المقطوع من المحور الصادي لشركة GE و α_2، α_3، α_4 الأجزاء المقطوعة من المحور الصادي للشركات GM، US و WEST عن الجزء المقطوع من المحور الصادي والخاص بشركة GE. باختصار فإن GE هي الشركة المقارنة. وبالطبع فإن لك مطلق الحرية في اختيار الشركة التي تكون شركة مقارنة وبشكل محدد إذا أردت قيمًا واضحة للأجزاء المقطوعة من المحور الصادي لكل شركة، فإنه يمكنك استخدام الأربعة متغيرات الوهمية السابق ذكرها، ثم نقوم بتحليل الانحدار من خلال نقطة الأصل، أو بمعنى آخر إسقاط الجزء المقطوع من المحور الصادي الموجود في (3.3.16) وإذا لم تقم بذلك فقد وقعت في مصيدة (مشكلة) المتغير الوهمي.

وبما أننا نستخدم أسلوب المتغيرات الوهمية لتقدير نموذج التأثيرات الثابتة، فقد عرف هذا النموذج تاريخيًا باسم نموذج المربعات الصغرى بالمتغير الوهمي (LSDV). وبالتالي فمصطلح التأثيرات الثابتة و LSDV يتم استخدامها تبادليًا. ولاحظ أن نموذج LSDV الموجود في (3.3.16) يعرف أيضًا باسم نموذج التغاير covaiance model، ويسمى X_2، X_3 بالمتغيرات المساعدة covaiate.

نتائج نموذج (3.3.16) هي كالتالي:

$$\hat{Y}_{it} = -245.7924 + 161.5722D_{2i} + 339.6328D_{3i} + 186.5666D_{4i} + 0.1079X_{2i} + 0.3461X_{3i}$$

$$\text{se} = (35.8112) \quad (46.4563) \quad (23.9863) \quad (31.5068) \quad (0.0175) \quad (0.0266)$$

$$t = (-6.8635) \quad (3.4779) \quad (14.1594) \quad (5.9214) \quad (6.1653) \quad (12.9821)$$

$$R^2 = 0.9345 \qquad d = 1.1076 \qquad \text{df} = 74 \qquad \textbf{(4.3.16)}$$

قارن هذا الانحدار مع (1.3.16). نجد أنه في (4.3.16) كل معاملات الانحدار المقدرة لها معنوية عالية فردية، حيث إن قيمة P-value لقيمة t المقدرة صغيرة جداً. الجزء المقطوع من المحور الصادي للشركات الأربع مختلف إحصائيًا، بحيث إن 245.7924- لشركة GE

و (161.5722 + 245.7924 - =) 84.220 - لشركة GM

و (339.6328 + 245.7924 - =) 93.8774 لشركة US

و (186.5666 + 245.7924 - =) 59.2258 - لشركة WEST. هذه الاختلافات قد تكون بسبب الصفات الخاصة والفريدة لكل شركة، مثل اختلاف أسلوب الإدارة أو الموهبة الإدارية. أي النموذجين أفضل - (1.3.16) أو (4.3.16)؟ الإجابة يجب أن تكون واضحة. باستخدام المعنوية الإحصائية للمعاملات المقدرة و زيادة قيمة R^2 وارتفاع قيمة إحصاء d (Durbin- Watson). نقترح أن نموذج (1.3.16) كان يعاني من مشكلة خطأ التوصيف. الزيادة في قيمة R^2 عمومًا لاتعتبر شيئًا مستغربًا بعد أن أدخلنا متغيرات جديدة في النموذج (4.3.16). من الممكن استخدام اختيار ما للنموذجين. بالمقارنة مع نموذج (4.3.16) فإن نموذج (1.3.16) يعتبر نموذجًا مقيداً، حيث إننا فرضنا جزءا مقطوعًا من المحور الصادي واحد فقط لكل الشركات. وبالتالي من الممكن استخدام اختيار F المقيد الذي تم شرحه في الفصل (8). باستخدام المعادلة (10.7.16) يمكن للقارئ أن يتأكد من أن قيمة F للمثال الحالي هي:

$$F\frac{(R^2_{\text{UR}} - R^2_R)/3}{(1 - R^2_{\text{UR}})/74} = \frac{(0.9345 - 0.7565)/3}{(1 - 0.9345)/74} = 66.9980 \qquad \textbf{(5.3.16)}$$

بحيث أن قيمة R^2 المقيدة حصلنا عليها من (1.3.16) أما قيمة R^2 غير المقيدة فحصلنا عليها من (16.3.4) وعدد القيود هي 3. حيث إن نموذج (1.3.16) يفترض أن الجزء المقطوع من المحور الصادي لكل من GE، GM، US و WEST متساوي. قيمة F المساوية لـ 66.9980 واضح أنها قيمة معنوية (درجات حرية البسط 3، ودرجات حرية المقام 74) وبالتالي الانحدار المقيد (1.3.16) يبدو غير معقول.

تأثير الزمن: كما سبق وذكرنا، فإن المتغيرات الوهمية تعبر عن تأثير اختلاف المشاهدة (الشركة) من الممكن إدخال تأثير الزمن، بمعنى أن دالة استثمار جرينفيلد Grunfeld تتغير مع الزمن بسبب متغيرات مثل التغيرات التكنولوچية، تغيير القواعد الحكومية أو سياسات الضرائب أو عوامل خارجية مثل الحروب أو الصراعات. من السهل اعتبار عامل التغير في الزمن عن طريق استخدام متغيرات وهمية للزمن، متغير وهمي واحد لكل سنة. وبما أن لدينا بيانات عن 20 سنة من عام 1935 إلى 1954. فمن الممكن استخدام 19 متغيراً وهميًا زمنيًا (لماذا؟) ونكتب النموذج كالتالي:

$$Y_{it} = \lambda_0 + \lambda_1\text{Dum35} + \lambda_2\text{Dum36} + \cdots + \lambda_{19}\text{Dum53} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \qquad \textbf{(6.3.16)}$$

بحيث إن DUM 35 يأخذ القيمة 1 للمشاهدات في العام 1935 و 0 بخلاف ذلك وهكذا. تم استخدام عام 1654 كعام الأساس، وبالتالي فإن الأجزاء المقطوعة من المحور الصادي معطاة بدلالة λ_0 (لماذا؟) لن نقوم باستعراض نتائج الانحدار الموجود في (6.3.16) حيث إنه لم يكن لأي من المعاملات المقدرة للزمن أي معنوية إحصائية وكانت قيمة R^2 الخاصة بـ (6.3.16) هي 0.7697 وهي تزيد عن قيمة R^2 الخاصة بـ (1.3.16) والتي كانت 0.7565 بمقدار 0.0132 . ومتروك كتمرين للقارئ أن يثبت أن هذه الزيادة، باستخدام اختيار F المقيد، زيادة غير معنوية، وذلك يجعلنا نعتقد أن العام أو التأثير الزمني غير معنوي، مما يعني أن دالة الاستثمار لم تتغير مع الزمن.

قد رأينا بالفعل أن تأثير اختلاف المشاهدات (الشركات) له معنوية إحصائية، ولكن اختلاف الزمن غير معنوي. هل من الممكن أن يكون نموذجنا الحالي يعاني من مشكلة خطأ التوصيف، حيث إنه لم يتم إدراك عامل التفاعل بين اختلاف الشركات واختلاف الزمن معًا؟ دعنا ندرس هذه الإمكانية.

3- معاملات الميل ثابتة، ولكن الجزء المقطوع من المحور الصادي يختلف باختلاف المشاهدة وباختلاف الزمن أيضاً :

3 – Slop coefficients constant but the intercept varies over individuals as well as time

لدراسة تلك الحالة، دعنا نمزج (4.3.16) مع (6.3.16) كالتالي :

$$Y_{it} = \alpha_1 + \alpha_2 D_{\text{GM}_i} + \alpha_3 D_{\text{US}_i} + \alpha_4 D_{\text{WEST}_i} + \lambda_0 + \lambda_1 \text{Dum35} + \cdots + \lambda_{19}\text{Dum53} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_{it} \quad (7.3.16)$$

عندما نقوم بهذا الانحدار، نجد أن كلاً من المتغيرات الوهمية المعبرة عن اختلاف الشركة ومعاملات المتغيرات X's لها معنوية إحصائية منفردة، ولكن لاتوجد أي معنوية لأي متغير وهمي خاص بالزمن، وبالتالي فقد عدنا إلى (4.3.16) وبالتالي الاستنتاج العام هو أن هناك تأثيراً للشركات، ولكن لايوجد تأثير لاختلاف الزمن. بمعنى آخر، فإن دالة الاستثمار للشركات الأربع واحدة باستثناء الجزء المقطوع من المحور الصادي. في كل الحالات، فقد وجدنا تأثيراً كبيراً للمتغيرات X على المتغير Y.

4- كل المعاملات تختلف باختلاف المشاهدات :

4 - All coefficients vary across individuals

هنا دعنا نفترض أن الجزء المقطوع من المحور الصادي، ومعاملات الميل تختلف باختلاف المشاهدات أو الوحدات المقطعية. وبالتالي، فإن دالة الاستثمار لكل من GE، GM، US و WEST مختلفة. من السهل جعل نموذج LSDV يشتمل على مثل هذه الحالة. بالعودة إلى نموذج (4.3.16) والذي قدمنا فيه المتغيرات الوهمية المعبرة عن الشركة بشكل إضافي. لكن في الفصل (9) الخاص بالمتغيرات الوهمية، أوضحنا كيف أن تفاعل المتغيرات الوهمية مع الميل ممكن أن يحدث فرقًا في معاملات الانحدار (الميل). للقيام بذلك في إطار دالة استثمار جرنفيلد Grunfeld كل ما نحتاج لعمله هو ضرب المتغيرات الوهمية للشركة في كل متغير من المتغيرات X [هذا سيضيف ستة متغيرات جديدة إلى (4.3.16)]، وبالتالي فنحن نقدر النموذج التالي :

$$Y_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \gamma_1(D_{2i}X_{2it}) + \gamma_2(D_{2i}X_{3it}) + \gamma_3(D_{3i}X_{2it}) + \gamma_4(D_{3i}X_{3it}) + \gamma_5(D_{4i}X_{2it}) + \gamma_6(D_{4i}X_{3it}) + u_{it} \quad (8.3.16)$$

نلاحظ أن الـ γ's هي معاملات الميل التفاضلية، كما أن α_2، α_3 و α_4 هي الجزء المقطوع من المحور الصادي التفاضلي. إذا كان واحد أو أكثر من المعاملات γ له معنوية إحصائية، فإنه من الممكن القول بأن واحداً أو أكثر من معاملات الميل مختلف عن المجموعة الأساسية.

على سبيل المثال، افترض أن β_2 و γ_1 لهما معنوية إحصائية. في مثل هذه الحالة، فإن $(\beta_2+\gamma_1)$ سيعطي القيمة الخاصة بمعامل ميل X_2 لشركة جنرال موتورز General Motors مما يعني أن معامل انحدار GM لـ X_2 مختلف عن نظيره لشركة جنرال إليكتريك General Electric وهي المجموعة الأساسية (المقارن بها).

إذا كانت كل معاملات الميل التفاضلية والأجزاء المقطوعة من المحور الصادي التفاضلية جميعها معنوية إحصائياً، فمن الممكن أن نستنتج أن دالة الاستثمار لكل من جنرال موتورز والولايات المتحدة للصلب ووستنجهوس تختلف عن جنرال إليكتريك. إذا كانت هذه هي الحالة بالفعل، فيكون هناك معنى من تقدير النموذج التجميعي (1.3.16).

دعنا ندرس نتائج انحدار (8.3.16). لسهولة القراءة، فإن نتائج الانحدار الخاصة بنموذج (8.3.16) معطاة في شكل جدولي في جدول (16.2) كما توضح هذه النتائج، فإن y على علاقة معنوية مع X_2، X_3. عموماً هناك بعض معاملات الانحدار التفاضلية لها معنوية إحصائية. فمثلاً، معامل انحدار X_2 المساوي 0.0902 لشركة GE ولكن نظيره لشركة GM يساوي (0.0902 + 0.0902) 0.1828 ولاتوجد أي معنوية إحصائية لأي من الأجزاء المقطوعة من المحور الصادي التفاضلية.

جدول (2.16) نتائج انحدار (8.3.16)

Variable	Coefficient	Std. error	*t* value	*p* value
Intercept	-9.9563	76.3518	-0.1304	0.8966
D_{2i}	-139.5104	109.2808	-1.2766	0.2061
D_{3i}	-40.1217	129.2343	-0.3104	0.7572
D_{4i}	9.3759	93.1172	0.1006	0.9201
X_{2i}	0.0926	0.0424	2.1844	0.0324
X_{3i}	0.1516	0.0625	2.4250	0.0180
$D_{2i}X_{2i}$	0.0926	0.0424	2.1844	0.0324
$D_{2i}X_{3i}$	0.2198	0.0682	3.2190	0.0020
$D_{3i}X_{2i}$	0.1448	0.0646	2.2409	0.0283
$D_{3i}X_{3i}$	0.2570	0.1204	2.1333	0.0365
$D_{4i}X_{2i}$	0.0265	0.1114	0.2384	0.8122
$D_{4i}X_{3i}$	-0.0600	0.3785	-0.1584	0.8745

$R^2 = 0.9511 \quad d = 1.0896$

من كل ماسبق، نخلص إلى أن دالة الاستثمار للشركات الأربع مختلفة. وبالتالي فإن البيانات الخاصة بالشركات الأربع غير قابلة للتجميع، بمعنى ضرورة تقدير دالة الاستثمار لكل شركة على حدة (انظر تمرين 13.16). وهذا يوضح أن انحدار البيانات الطولية قد يكون غير مناسب في بعض الحالات، بغض النظر عن بيانات السلاسل الزمنية والبيانات المقطعية المتاحة.

لابد من ذكر تحذير يجب وضعه في الاعتبار عند استخدام نماذج التأثير الثابت أو LSDV. فعلى الرغم من سهولة استخدام نماذج LSDV ، فإن هناك بعض المشاكل التي يجب وضعها في الاعتبار عند الاستخدام.

أولاً : إذا استخدمنا عدداً كبيراً من المتغيرات الوهمية كما هو الحال في نموذج (7.3.16) سنجد مشكلة في درجات الحرية. ففي حالة نموذج (7.3.16) لدينا 80 مشاهدة ولكن درجات الحرية هي 55 فقط. فقد فقدنا 3 درجات لثلاثة متغيرات وهمية للشركات و 19 درجة حرية مفقودة للمتغيرات الوهمية الخاصة بالعام (الزمن) و 2 درجة لمعاملات الميل، و 1 درجة للجزء العام المقطوع من المحور الصادي.

ثانيًا، بوجود متغيرات عديدة في النموذج، فهناك دائمًا احتمال ظهور مشكلة الارتباط الخطي المتعدد، والذي يحدث صعوبة في تقدير واحد أو أكثر من معلمات النموذج.

ثالثًا، افترض أنه في FEM (1.3.16) تم إضافة متغيرات مثل النوع، لون البشرة والعرق، وهي متغيرات ثابتة بالنسبة للزمن، حيث إن نوع المشاهدة أو لونها أو عرقها لن يتغير بتغير الزمن. وبالتالي فإن طريقة LSDV قد تكون غير قادرة على توضيح تأثير مثل هذه العوامل غير المتغيرة مع الزمن.

رابعًا، لابد من التفكير بعمق في مقدار الخطأ u_{it}. كل النتائج التي قدمناها حتى الآن معتمدة على افتراض أن مقدار الخطأ يتبع الفروض التقليدية الخاصة بالتوزيع الطبيعي $u_{it} \sim N(0, \sigma^2)$. وبما أن الترميز i يعبر عن المشاهدة المقطعية و t، يعبر عن مشاهدة السلسلة الزمنية، فإن الفروض التقليدية لـ u_{it} لابد من تعديلها. وهناك العديد من الأشياء الممكنة هنا وهي:

1 - من الممكن افتراض أن تباين الخطأ واحد بالنسبة المفردات المقطعية أو بمعنى أخر من الممكن افتراض ثبات التباين.

2 - بالنسبة لكل مشاهدة من الممكن افتراض عدم وجود ارتباط ذاتي مع الزمن. وبالتالي فعلى سبيل المثال، من الممكن افتراض أن مقدار الخطأ في دالة استثمار شركة جنرال موتورز لايوجد فيه ارتباط ذاتي. أو من الممكن افتراض أنه مرتبط خطيًا مثلاً من النوع AR(1).

3 - بمعلومية عام معين، فإنه من الممكن لمقدار الخطأ الخاص بشركة جنرال موتورز أن يكون مرتبطًا مع مقدار الخطأ مثلاً الخاص بـ U.S.Steel أو كلاً من US.Steel و Westing house[7]. أو من الممكن افتراض عدم وجود مثل هذا الارتباط.

4 - من الممكن التفكير في توليفات واحتمالات عديدة خاصة بمقدار الخطأ، وستدرك سريعًا أن السماح بواحد أو أكثر من هذه الحالات، سيجعل التحليل أكثر تعقيدًا. الوقت والتعقيد الرياضي يجعلان من الصعب التعمق في دراسة مثل هذه الحالات.

لقراءة ممكنة عن هذه الاحتمالات ارجع إلى Dielman and Kmenta[8]. عمومًا بعض هذه المشاكل من الممكن تجنبها إذا استخدمت النموذج المسمى نموذج التأثيرات العشوائية والذي سنناقشه في الفقرة التالية.

4.16 تقدير نماذج انحدار البيانات الطولية.. طريقة التأثيرات العشوائية: ESTIMATION OF PANEL DATA REGRESSION MODELS: THE RANDOM EFFECTS APPROACH

على الرغم من سهولة تطبيق طريقة التأثيرات الثابتة أو LSDV، فإن المقابل يكون باهظًا بالنظر إلى درجات الحرية إذا كان لدينا العديد من الوحدات المقطعية، بالإضافة إلى ذلك، فإن Kmenta لاحظ التالي:

هناك سؤال واضح خاص بنموذج التغاير (LSDV) ألا وهو: هل المتغيرات الوهمية وما يتبعها من خسارة في درجات الحرية تعتبر فعلاً مهمة؟ وهذا التساؤل حول نموذج التغاير يأتي من حقيقة أننا لجأنا إلى المتغيرات الوهمية، لأننا فشلنا في

(7) هذا يؤول إلى ما يسمى نموذج الانحدار غير المرتبط ظاهريًا (SURE) والمقترح أساسًا من Arnold Zellner. للدراسة انظر في Terry E.Dielman, op., cit.

(8) Dielman, op. cit., Sayrs, op.cit., Jan Kmenta, Elements of Econometrics, 2nd ed., Macmillan, New York, 1986, chap.12.

إيجاد متغيرات مفسرة مناسبة للنموذج لاتتغير مع الزمن (أو محتمل وجود بعض المتغيرات المفسرة التي تتغير مع الزمن، ولكن تأخذ نفس القيم لكل الوحدات المقطعية) وبالتالي وجود المتغيرات الوهمية ماهو إلا علاج لهذه المشكلة.(9)

إذا كانت المتغيرات الوهمية هي في الحقيقة تمثيل لانعدام المعرفة بالنموذج الحقيقي، لماذا لايتم التعبير عن ذلك من خلال مقدار الخطأ u_{it}؟ هذا هو بالفعل الأسلوب أو الطريقة المقترحة المسماة نموذج عناصر الخطأ (ECM)، أو نموذج التأثيرات العشوائية (REM). الفكرة الرئيسية تبدأ بالمعادلة (2.3.16):

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \qquad (1.4.16)$$

بدلاً من التعامل مع β_{1i} على أنها ثابتة، يتم افتراض أنها متغير عشوائي له توقع β_1 (بدون ترميز i)، والقيمة المقطوعة من المحور الصادي لشركة ما من الممكن التعبير عنه كالتالي:

$$\beta_{1i} = \beta_1 + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (2.4.16)$$

حيث إن ε_i هو مقدار الخطأ العشوائي، وله توقع يساوي صفراً، وتباين يساوي σ_ε^2. ما نعنيه من ذلك هو أن الأربع شركات الموجودة في العينة المسحوبة من مجتمع أكبر تحتوي على كل الشركات المماثلة ولها جميعًا توقع مشترك للجزء المقطوع من المحور الصادي يساوي β_1، والفروق بين الوحدات في قيم الجزء المقطوع من المحور الصادي الخاص بكل شركة يتم تمثيله في مقدار الخطأ ε_i.

بالتعويض عن (2.4.16) في (1.4.16) نحصل على:

$$\begin{aligned} Y_{it} &= \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \varepsilon_i + u_{it} \qquad (3.4.16) \\ &= \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + w_{it} \end{aligned}$$

بحيث إن:

$$w_{it} = \varepsilon_i + u_{it} \qquad (4.4.16)$$

(9) Kmenta, op.cit., p.633:

مـقـدار الخطأ المركب w_{it} مكون من جـزءين، ε_i وهو يمثل عنصـر الخطأ الخاص بالبـيـان المقطعي أو بتـحـديد المفـردات والمقـدار u_{it} وهو عنصـر الخطأ الناتج من دمج السلاسل الزمنية مع البيانات المقطعية. وبالتالي، فإن تسمية النموذج بنموذج عناصر الخطأ يأتي أساسًا من مقدار الخطأ المركب w_{it} والمكون من جزءين (أو أكثر) من عناصر الخطأ.

الفروض التقليدية الموجودة في ECM هي:

$$
\begin{aligned}
\varepsilon_i &\sim N(0, \sigma_\varepsilon^2) \\
u_{it} &\sim N(0, \sigma_u^2) \qquad\qquad (5.4.16) \\
E(\varepsilon_i u_{it}) = 0 &\qquad E(\varepsilon_i \varepsilon_j) = 0 \quad (i \neq j) \\
E(u_{it}u_{is}) = E(u_{it}u_{jt}) &= E(u_{it}u_{js}) = 0 \quad (i \neq j; t \neq s).
\end{aligned}
$$

وبالتالي، فعناصر الخطأ غير مرتبطة مع بعضها البعض، ولايوجد ارتباط ذاتي بالنسبة لوحدات البيانات المقطعية أو لوحدات السلاسل الزمنية.

لاحظ أن هناك فرقًا بين FEM و ECM. في الـ FEM كل وحدة بيان مقطعي لها الجزء المقطوع من المحور الصادي (الثابت) الخاص بها لكل N من هذه القيم من الوحدات المقطعية. في الـ ECM بخلاف ذلك فإن الجزء المقطوع من المحور الصادي β_1 يمثل القيمة المتوقعة لكل بيان مقطعي ومقدار الخطأ ε_i يمثل الانحراف (العشوائي) للأجزاء المقطوعة من المحور الصادي للمفردات عن القيمة المتوقعة. عمومًا ضع في الاعتبار أن ε_i لاتلاحظ مباشرة، بل هي ما يطلق عليه المتغير غير المشاهد أو الخفي.

كنتيجة للفروض الموضحة في (5.4.16) نحصل على التالي:

$$E(w_{it}) = 0 \qquad (6.4.16)$$

$$\operatorname{var}(w_{it}) = \sigma_\varepsilon^2 + \sigma_u^2 \qquad (7.4.16)$$

والآن إذا كـان $\sigma_\varepsilon^2 = 0$ بمعنى أنه لاتوجـد فـروق بين نموذج (1.2.16) ونموذج (3.4.16) ففي هذه الحالة يمكن ببساطة تجميع كل المشاهدات (المقطعية والسلاسل الزمنية) ونقوم بإجراء انحدار تجميعي كما فعلنا في (1.3.16). كما نرى من (7.4.16) مـقـدار الخطأ w_{it} له تباين ثابت، وعـمـومًا من الممكن إثبات أن w_{it} و w_{is} $(t \neq s)$

مرتبطان، بمعنى أن مقدار الخطأ لوحدة من الوحدات المقطعية عند نقطتين مختلفتين في الزمن مرتبطتان. معامل الارتباط $\text{corr}(w_{it}, w_{is})$ هو كالتالي:

$$\text{corr}(w_{it}, w_{is}) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2} \quad (8.4.16)$$

لاحظ صفتين رئيسيتين في معامل الارتباط السابق. **أولاً**، لأي قيمة محددة لوحدة بيان مقطعي فإن قيمة الارتباط بين مقادير الأخطاء عند قيمتين مختلفتين في الزمن تظل كما هي، بغض النظر عن مدى البعد الزمني بين القيمتين وهذا واضح من (8.4.16). هذا مخالف تمامًا كما هو معروف في أسلوب السلاسل الزمنية من الدرجة الأولى AR(1) المشروح في الفصل (12)، والذي لاحظنا من خلاله أن الارتباط بين النقاط الزمنية يقل بمرور الزمن. **ثانيًا**، شكل الارتباط المعطى في (8.4.16) يبقى كما هو بالنسبة لجميع الوحدات المقطعية.

إذا لم نضع في الاعتبار هذا الارتباط، وقمنا بتقدير (3.4.16) بطريقة OLS ستكون النتائج الخاصة بالمقدرات غير كفء. أفضل النماذج المناسبة هنا هو أسلوب المربعات الصغرى العام (GLS) لن نناقش الخلفية الرياضية لطريقة GLS في هذا الكتاب نظرًا لصعوبتها النسبية[10]. وبما أن معظم حزم البرامج الإحصائية الحديثة تحتوي على أكواد خاصة لتقديرات ECM (وأيضًا FEM) سنقوم فقط بعرض مثل هذه النتائج والخاصة بمثال الاستثمار. ولكن قبل أن نقوم بذلك، فإننا نرغب في الإشارة إلى أنه يمكن تطبيق (4.4.16) بحيث يسمح بوجود مقدار خطأ عشوائي ممثل للتغير في الزمن (انظر تمرين 6.16).

نتائج تقديرات ECM لدالة استثمار Grunfeld معطاة في جدول (3.16) بعض النقاط يجب ملاحظتها في هذا الانحدار. أولاً، إذا جمعنا قيم التأثيرات العشوائية المعطاة للأربع شركات سنحصل على صفر، أو يجب أن نحصل على صفر (لماذا؟). ثانيًا، القيمة المتوقعة لمقدار الخطأ العشوائي ε_i هي الجزء المقطوع من المحور الصادي ذي القيمة 73.0353−. قيمة التأثير العشوائي لـ GE تساوي 169.9283− وهي عبارة عن الفرق بين مقدار الخطأ العشوائي لـ GE عن القيمة المشتركة للجزء المقطوع من المحور الصادي. تفسيرات مماثلة يمكن الحصول عليها للثلاث قيم الأخرى للتأثيرات العشوائية. ثالثًا، قيمة R^2 تم الحصول عليها من انحدار GLS المحول.

(10) القارئ المهتم بذلك يمكنه الرجوع إلى Kmenta, op.cit., pp. 625-630 سيجد مناقشة وافية للموضوع.

إذا قارنا النتائج الخاصة بنموذج ECM المعطاة في جدول (3.16) مع نظيرها الخاص بنموذج FEM، ستجد أنه بوجه عام قيم المعاملات الخاصة بالمتغيرين المفسرين لاتختلف بشكل كبير باستثناء القيم المعطاة في جدول (2.16) عندما جعلنا معاملات الميل لهذين المتغيرين تختلف باختلاف الوحدة المقطعية.

جدول (3.16) تقديرات ECM لدالة استثمار Grunfeld

Variable	Coefficient	Std. error	t statistic	p value
Intercept	-73.0353	83.9495	-0.8699	0.3870
X_2	0.1076	0.0168	6.4016	0.0000
X_3	0.3457	0.0168	13.0235	0.0000
Random effect:				
GE	-169.9282			
GM	-9.5078			
USS	165.5613			
Westinghouse	13.87475			
$R^2 = 0.9323$ (GLS)				

5.16 نماذج التأثيرات الثابتة (LSDV) ومقارنتها مع نماذج التأثيرات العشوائية: FIXED EFFECTS (LSDV) VERSUS RANDOM EFFECTS MODEL

السؤال المهم لأي باحث هو: أي النموذجين أفضل، FEM أم ECM؟ إجابة هذا السؤال تكمن وراء الفروض التي يضعها الباحث عن الارتباط المحتمل بين المفردات أو الوحدات المقطعية ومقدار الخطأ ε_i والمتغيرات المنحدرة X.

إذا كان هناك افتراض بأن ε_i و X's غير مرتبطين، فإن نموذج ECM يكون أفضل. في حين أنه إذا كان ε_i و X's مرتبطين فإن FEM يكون أفضل. لماذا قد يتوقع الباحث وجود ارتباط بين عناصر الخطأ ε_i وواحد أو أكثر من المتغيرات المنحدرة؟ دعنا نستعرض المثال التالي. افترض أن لدينا عينة عشوائية من أعداد كثيرة من المفردات ونريد نمذجة دالة الأجور الخاصة بها (المكاسب).

افترض أن مكاسب الشخص دالة في التعليم، خبرة العمل وإلى ما غير ذلك من عوامل مفسرة أخرى. الآن دعنا نفترض ε_i معبر عن القدرات الفطرية والخلفية الأسرية للشخص وهكذا، وبالتالي عندما نحاول نمذجة دالة الأجر أو المكسب والتي تشتمل على المقدار ε_i فإنه من المحتمل أن يكون مرتبطًا مع التعليم، فالقدرات الفطرية

والخلفية الأسرية على علاقة وثيقة مع مستوى تعليم الفرد. كما هو مذكور في Wooldridge contends فإن: "في العديد من التطبيقات، يكون السبب الرئيسي لاستخدام البيانات الطولية هو السماح لمقدار التأثير غير الملاحظ (ε_i) أن يكون مرتبطًا مع المتغيرات المفسرة[11].

الفروض الخاصة بنموذج ECM معتمدة على أن المقدار ε_i مسحوب عشوائيًا من مجتمع أكبر. ولكن أحيانًا يكون الأمر بخلاف ذلك، فمثلاً افترض أننا نريد دراسة معدل الجريمة في الولايات الخمسين بالولايات المتحدة الأمريكية كما هو واضح في مثل هذه الحالة الافتراض بأن الـ 50 ولاية هي عينة عشوائية غير قابلة للاستخدام.

مع الوضع في الاعتبار كل هذه الأمور السابقة. ماذا أيضًا من الممكن الاعتماد عليه للاختيار ما بين FEM و ECM؟ دعنا نستعرض ما قام به Judge et al في هذا الشأن وهو كالتالي[12]:

1 - إذا كان T (عدد بيانات السلاسل الزمنية) كبيرًا و N (عدد الوحدات المقطعية) سيكون من المحتمل وجود فروق بسيطة بين قيم المقدرات التي نحصل عليها من FEM و ECM. وبالتالي الاختيار بين النموذجين سيرجع إلى اختيار الأسهل حسابيًا. ومن هذا المنطلق فإن FEM قد يكون أفضل.

2 - عندما تكون N كبيرة و T صغيرة، التقديرات التي سنحصل عليها من النموذجين ستكون بينها فروق معنوية. تذكر أنه في نموذج ECM $\beta_{1i}=\beta_1+\varepsilon_i$ بحيث إن ε_i هو مقدار عشوائي للبيانات المقطعية في حين أنه في نموذج FEM نعامل β_{1i} على أنه ثابت وليس عشوائيًا. في هذه الحالة الأخيرة، فإن الاستدلال الإحصائي مشروط على الوحدات المقطعية الموجودة في العينة. هذا يكون مناسبًا إذا اقتنعنا بشدة أن المفردات أو الوحدات المقطعية في العينة ليست مسحوبة عشوائيًا من مجتمع أكبر. في هذه الحالة فإن FEM يكون أفضل. عمومًا إذا تم اعتبار الوحدات المقطعية في العينة مسحوبة كعينة عشوائية من مجتمع أكبر، فإن ECM يكون أفضل، وفي هذه الحالة يكون الاستدلال الإحصائي غير مشروط.

(11)Wiildridges, op.cit., p.450.

(12) Judge et al., op.cit., pp. 489-491.

3 - إذا كان هناك ارتباط بين مقدار الخطأ ε_i وواحد أو أكثر من المتغيرات المنحدرة، فإن مقدرات ECM تكون متحيزة. في حين المقدرات التي نحصل عليها من FEM تكون غير متحيزة.

4 - إذا كانت N كبيرة و T صغيرة، وكانت الفروض الخاصة بـ ECM متحققة، فإن مقدرات ECM تكون أكثر كفاءة من مقدرات FEM[13].

هل هناك اختبار معروف يساعدنا على الاختيار بين FEM و ECM؟

نعم هذا الاختيار قام به Housman في عام 1978[14]. لن نقوم بمناقشة تفاصيل هذا الاختيار، حيث إن ذلك خارج نطاق هذا الكتاب[15].

الفرض العدمي في اختبار Housman هو أن مقدرات FEM ومقدرات ECM غير مختلفة. قيمة إحصاء الاختبار الذي قام به Housman تتبع تقاربيًا توزيع X^2. إذا تم رفض الفرض العدمي يكون ذلك دليلاً على أن ECM غير مناسب، ومن الأفضل استخدام FEM، ويكون في هذه الحالة الاستدلال الإحصائي مشروطًا على ε_i الموجود في العينة.

إلى جانب اختبار Housman يجب أن نضع في الاعتبار التحذيرات التي ذكرها DiNardo و Johnston. ففي مسألة الاختيار بين نماذج التأثيرات الثابتة أو نماذج التأثيرات العشوائية فقد أوضحا أن: ". . لاتوجد قاعدة بسيطة تساعد الباحث في الاختيار بين التأثيرات الثابتة و Charybdis لقياسات الأخطاء والاختيارات المتحركة. وعلى الرغم من أن البيانات الطولية تمثل تقدمًا عن البيانات المقطعية فإنها لاتعطي حلولاً لكل مشاكل الاقتصاد القياسي"[16].

---

(13) Taylor أوضح أنه إذا كان $T \geq 3$ و $(N - K) \geq 9$ بحيث إن K هو عدد المتغيرات المنحدرة، فإن الفرضية تتحقق. انظر:

(14)W.E. Taylor, "Small sample consideration. In Estimating from panel data", Journal of Econometrics, vol. 13, 1980, pp.203- 223.

(15) J.A. Housman, "Specification in tests in Econometrics, vol. 46, 1978, pp. 1251- (27)

(15) لمزيد من التفاصيل. انظر Baltagi, op.cit., pp. 68-73.

(16) Jack Johnson and Jonn Di Nardo, Econometric Methods, 4th ed., Mc Craw- Hill, 1997, p. 403.

6.16 انحدار البيانات الطولية.. بعض التعليقات الإستنتاجية :
PANEL DATA REGRESSIONS: SOME CONCLUDING COMMENTS

كما سبق الذكر، فإن موضوع البيانات الطولية واسع ومعقد. فإننا فقط تعرضنا لسطح الموضوع. من بين النقاط التي لم نتعرض لها، **دعنا نذكر التالي** :

1 - اختبارات الفروض للبيانات الطولية .

2 - اختلاف التباين والارتباط الذاتي في ECM .

3 - البيانات الطولية غير المتوازنة .

4 - نماذج البيانات المتحركة، عندما تكون القيم الزمنية للمتغير المنحدر عليه (Y_{it}) تظهر كمتغير مفسر .

5 - المعادلات الآنية الخاصة بالبيانات الطولية .

6 - المتغيرات التابعة النوعية والبيانات الطولية .

واحد أو أكثر من هذا المواضيع، يمكن القراءة عنها بالتفصيل في المراجع المذكورة بالفصل، وعلى الباحث اللجوء إلى مثل هذه المراجع للتعرف أكثر على هذه المواضيع. في هذه المراجع، هناك أيضًا العديد من الدراسات التجريبية في مجالات عديدة من الاقتصاد وإدارة الأعمال، والتي استخدمت نماذج الانحدار الخاصة ببيانات طولية .

المبتدئين في هذا المجال، من الأفضل لهم قراءة بعض هذه التطبيقات لمعرفة كيف يتم فعليًا التعامل مع مثل هذه النماذج .

7.16 التلخيص والنتائج : SUMMARY AND CONCLUSIONS

1 - نماذج الانحدار الطولي معتمدة على بيانات طولية. البيانات الطولية تتكون من مشاهدات خاصة بنفس المفردات أو وحدات مقطعية عبر فترة زمنية معينة .

2 - هناك العديد من مميزات استخدام البيانات الطولية. أولاً: زيادة حجم العينة بشكل ملحوظ. ثانيًا: بدراسة المشاهدات المقطعية المكررة البيانات الطولية تجعل لدينا القدرة على دراسة نماذج لمتغيرات أكثر تعقيداً .

3 - على الرغم من مميزات البيانات الطولية فهناك العديد من مشاكل التقدير والاستدلال الإحصائي المرتبطة بها فيما أن هذه البيانات لها بعد خاص بالبيانات المقطعية وأخرى بالسلاسل الزمنية، فإن المشاكل المرتبطة بكل من البيانات المقطعية (مثل الارتباط الذاتي) لابد من وضعها في الاعتبار. هناك مشاكل أخرى إضافية مثل الارتباط بين المفردات أو الوحدات المقطعية عند نفس النقطة الزمنية.

4 - هناك العديد من الأساليب المستخدمة في التقدير، والتي تضع في الاعتبار واحداً أو أكثر من هذه المشاكل السابق ذكرها. النموذجان الأكثر شهرة هنا (1) نماذج التأثيرات الثابتة (FEM) و (2) نماذج التأثيرات العشوائية (REM)، أو نموذج عناصر الأخطاء (ECM).

5 - في FEM فإن الجزء المقطوع من المحور الصادي مسموح له بالاختلاف بين المفردات للتعبير عن حقيقة أن كل مفردة أو وحدة مقطعية لها صفات خاصة بها تجعلها مختلفة مع الآخرين. ولأخذ هذه الخاصية في الاعتبار، من الممكن أن يستخدم الفرد أسلوب المتغيرات الوهمية. الـ FEM التي تستخدم المتغيرات الوهمية معروفة باسم نموذج المربعات الصغرى ذي المتغير الوهمي FEM. (LSDV) مناسب أيضًا في الحالات التي يكون فيها الجزء المقطوع من المحور الصادي الخاص بالمفردات مرتبطًا مع واحد أو أكثر من المتغيرات المنحدرة.

من عيوب طريقة LSDV أنها تستهلك الكثير من درجات الحرية خصوصًا عندما يكون N، عدد الوحدات المقطعية كبير جداً. حيث سيكون لدينا في هذه الحالة N من المتغيرات الوهمية (بالإضافة إلى الجزء المقطوع من المحور الصادي المشترك).

6 - كبديل عن FEM لدينا ECM. في ECM يفترض أن الجزء المقطوع من المحور الصادي هو عينة عشوائية مسحوبة من مجتمع أكبر لها قيمة متوقعة ثابتة. وبالتالي الجزء المقطوع من المحور الصادي الخاص بكل مفردة يتم التعبير عنه بالفرق بينه وبين هذه القيمة المتوقعة الثابتة. إحدى مميزات الـ ECM عن الـ FEM هي الاقتصاد في درجات الحرية، حيث إننا لانحتاج لتقدير N من الأجزاء المقطوعة من المحور الصادي والخاصة بكل وحدة من الوحدات المقطعية نحن نحتاج فقط إلى تقدير القيمة المتوقعة للجزء المقطوع من المحور الصادي وتباينه. ECM هو الأفضل في المواقف التي يكون فيها الجزء المقطوع من المحور الصادي (العشوائي) لكل وحدة من الوحدات المقطعية غير مرتبط مع المتغيرات المنحدرة.

7 – اختيار Housman من الممكن استخدامه للاختيار بين FEM و ECM.

8 – على الرغم من زيادة استخدام البيانات الطولية في الأبحاث التطبيقية، وعلى الرغم من زيادة المتاح من مثل هذا النوع من البيانات، فإن انحدار البيانات الطولية قد لا يكون مناسبًا لجميع الحالات، لابد للباحث أن يستخدم أحكامًا عملية تتغير مع تغير الحالة التي تتم دراستها.

تمارين : EXERCISES

أسئلة Questions

1.16 ما هي الخصائص المميزة لكل من (a) بيانات المقطعية (b) بيانات السلاسل الزمنية و (c) بيانات الطولية؟

2.16 ما هو المقصود بنموذج التأثيرات الثابتة (FEM)؟ بما أن بيانات الطولية لها بعدان، بعد زمني وبعد فراغي، كيف تسمح FEM بالتعامل مع هذين البعدين؟

3.16 ما هو المقصود بنموذج عناصر الخطأ (ECM)؟ وكيف يختلف عن FEM؟ متى يكون ECM مناسبًا؟ ومتى يكون FEM مناسبًا؟

4.16 هل هناك فرق بين نموذج FEM، نموذج المربعات الصغرى بالمتغيرات الوهمية (LSDV) ونموذج التغاير؟

5.16 متى تكون نماذج الانحدار البيانات panel غير مناسبة؟ اعط أمثلة.

6.16 كيف يمكن تطبيق نموذج (4.4.16) بحيث يسمح بمقدار خطأ زمني؟ في مثل هذه الحالة ماذا سيحدث للمعادلات (6.3.16) و (7.3.16) و (8.3.16)؟

7.16 بالرجوع إلى المثال الخاص بإنتاج البيض وأسعاره المعطى في جدول (1.1) ما هو النموذج الذي قد يكون مناسبًا في مثل هذه الحالة، FEM أم ECM؟ ولماذا؟

8.16 في نتائج الانحدار المعطاة في (4.3.16) ما هو الجزء المقطوع من المحور الصادي الثابت لكل من الشركات الأربع؟ هل هذه التأثيرات مختلفة إحصائيًا؟

9.16 في مثال الاستثمار الذي تم استعراضه في الفصل، جدول (3.16) يعطي النتائج باستخدام ECM. لو قارنت هذه النتائج مع نظيرها المعطى في (4.3.16) ما هي النتائج العامة التي يمكن أن تستخلصها؟

10.16 بالاعتماد على دراسة عن التغيرات الديناميكية للدخل في ولاية ميتشجن، حاول Housman تقدير الأجر أو الكسب من خلال نموذج استخدم فيه عينة من 629 متخرجًا من المدارس الثانوية تم تتبعهم لفترة 6 سنوات وبالتالي إجمالي المشاهدات هو 3774 مشاهدة. المتغير التابع في هذه الدراسة هو لوغاريتم الأجر، أما المتغيرات المفسرة فهي العمر (مقسم إلى عدة مجموعات عمرية)، البطالة في السنوات السابقة، الصحة المتدنية في السنوات السابقة، العمل الحر، محل الإقامة (الجنوب= 1، 0 بخلاف ذلك)، طبيعة محل الإقامة (الريف= 1، 0 بخلاف ذلك). Housman استخدم كلاً من FEM و ECM. النتائج معطاة في جدول (4.16) (الأخطاء القياسية بين الأقواس).

(a) هل هذه النتائج لها معنى اقتصادي؟

(b) هل هناك فروق كبيرة في النتائج بين النموذجين؟ وإذا كان كذلك، ما هو السبب في هذه الفروق؟

(c) على أساس البيانات المعطاة في الجدول، ما هو النموذج (إذا كان هناك نموذج) الذي تختاره؟

جدول (4.16) معادلات الأجور (المتغير التابع: لوغاريتم الأجر)(*)

Variable	Fixed effects		Random effects	
1. Age 1 (20–35)	0.0557	(0.0042)	0.0393	(0.0033)
2. Age 2 (35–45)	0.0351	(0.0051)	0.0092	(0.0036)
3. Age 3 (45–55)	0.0209	(0.0055)	−0.0007	(0.0042)
4. Age 4 (55–65)	0.0209	(0.0078)	−0.0097	(0.0060)
5. Age 5 (65–)	−0.0171	(0.0155)	−0.0423	(0.0121)
6. Unemployed previous year	−0.0042	(0.0153)	−0.0277	(0.0151)
7. Poor health previous year	−0.0204	(0.0221)	−0.0250	(0.0215)
8. Self-employment	−0.2190	(0.0297)	−0.2670	(0.0263)
9. South	−0.1569	(0.0656)	−0.0324	(0.0333)
10. Rural	−0.0101	(0.0317)	−0.1215	(0.0237)
11. Constant	—	—	0.8499	(0.0433)
s^2	0.0567		0.0694	
Degrees of freedom	3,135		3,763	

(*) 3774 مفردة والأخطاء القياسية معطاة بين الأقواس.

Reproduced from cheng Hsiao, Analysis of Pauel Data, Cambridgo press, 1986, p. 42. Original source: J.A.Housman, "Specification tests in econometris", econometrica, vol. 46, 1978, pp. 1251- 1271.

مسائل : Problems

11.16 بالعودة إلى البيانات الموجودة في جدول (1.1)

(**a**) دع Y = المنتج من البيض (بالمليون) و X = سعر البيض (cents لكل دستة).
قدر النموذج: $Y_i = \beta_1 + \beta_2 X_i + u_i$ للسنوات 1990 و 1991 منفصلين.

(**b**) ادمج المشاهدات للعامين السابقين، وقدر نموذج الانحدار التجميعي، ما هي الفروض اللازمة لدمج هذه البيانات؟

(**c**) استخدم نموذج التأثيرات الثابتة، فرق بين العامين، واعرض نتائج الانحدار.

(**d**) هل من الممكن استخدام نموذج التأثيرات السابقة مع التفريق بين الـ 50 ولاية؟ علل إجابتك.

(**e**) هل من الأهمية التفرقة بين تأثير الولاية والتأثير العام؟ وإذا كان ذلك صحيحًا، كم عدد المتغيرات الوهمية اللازمة لفعل ذلك.

(**f**) هل نموذج عناصر الخطأ مناسب للاستخدام كنموذج لإنتاج البيض؟ علل إجابتك. إدرس كيف يمكنك تقدير مثل هذا النموذج مستخدمًا على سبيل المثال Eviews.

12.16 بالعودة إلى تمرين (11.16). قبل استخدام الانحدار التجميعي، لابد أن ترى ما إذا كانت البيانات قابلة للتجميع أم لا. ولهذا السبب قررت استخدام اختبار chow الذي تمت دراسته في الفصل (8). وضح الحسابات الأساسية اللازمة، ووضح إذا كان الانحدار التجميعي مناسبًا أم لا.

13.16 بالعودة إلى دالة استثمار Grunfeld التي تمت مناقشتها في فقرة 2.16.

(**a**) قدر دالة استثمار Grunfeld لكل من GE، GM، U.S. Steel و westinghouse لكل منها منفردة. نتائج دمج كل الـ 80 مشاهدة معطاة في (1.3.16).

(**b**) لتحديد ما إذا كان الانحدار التجميعي (1.3.16) مناسبًا أم لا، استخدم اختبار chow الذي سبق دراسته في الفصل (8).

ملاحظة :

احصل على RSS من الانحدار التجميعي، واحصل على RSS من كل من دوال الاستثمار الأربع، وبعد ذلك طبق اختيار chow.

(c) من اختيار chow، ما هي النتائج التي خلصت بها؟ إذا كانت النتيجة عدم استخدام الانحدار التجميعي، ماذا إذن يمكن أن يقال عن أهمية أسلوب انحدار البيانات الطولية؟

14.16 جدول (5.16) يعطي بيانات عن معدلات البطالة المدنية Y (%) وساعات التعويض المصنعية بالدولار الأمريكي X (مؤشر، 100 = 1992) لكندا، المملكة المتحدة والولايات المتحدة في الفترة 1980–1999.

اعتبر النموذج التالي :

$$Y_{it} = \beta_1 + \beta_2 X_{it} + u_{it} \qquad (1)$$

(a) مبدئيًا: ماذا تتوقع عن العلاقة بين X، Y؟ لماذا؟

(b) قدر النموذج المعطى في (1) لكل بلد.

(c) قدر النموذج بعد الدمج (تجميع) الـ 60 مشاهدة.

(d) قدر نموذج التأثيرات الثابتة.

(e) قدر نموذج عناصر الخطأ.

(f) ما هو النموذج الأفضل، FEM أم ECM؟ علل إجابتك.

جدول (5.16) معدل البطالة والساعات التعويضية المصنعية في الولايات المتحدة وكندا والمملكة المتحدة، 1980–1999

	United States		Canada		United Kingdom	
Observation	Compensation, $/hour	Unemployment, %	Compensation, $/hour	Unemployment, %	Compensation, $/hour	Unemployment, %
1980	55.6	7.1	49.0	7.2	43.7	7.0
1981	61.1	7.6	54.1	7.3	44.1	10.5
1982	67.0	9.7	59.6	10.6	42.2	11.3
1983	68.8	9.6	63.9	11.5	39.0	11.8
1984	71.2	7.5	64.3	10.9	37.2	11.7
1985	75.1	7.2	63.5	10.2	39.0	11.2
1986	78.5	7.0	63.3	9.2	47.8	11.2
1987	80.7	6.2	68.0	8.4	60.2	10.3
1988	84.0	5.5	76.0	7.3	68.3	8.6
1989	86.6	5.3	84.1	7.0	67.7	7.2
1990	90.8	5.6	91.5	7.7	81.7	6.9
1991	95.6	6.8	100.1	9.8	90.5	8.8
1992	100.0	7.5	100.0	10.6	100.0	10.1
1993	102.7	6.9	95.5	10.7	88.7	10.5
1994	105.6	6.1	91.7	9.4	92.3	9.7
1995	107.9	5.6	93.3	8.5	95.9	8.7
1996	109.3	5.4	93.1	8.7	95.6	8.2
1997	111.4	4.9	94.4	8.2	103.3	7.0
1998	117.3	4.5	90.6	7.5	109.8	6.3
1999	123.2	4.0	91.9	5.7	112.2	6.1

الساعات التعويضية بالدولار الأمريكي، مؤشر 100= 1992

Source: Economic report of the president, January 2001, table B 109, p. 399.

الفصل السابع عشر

نماذج الاقتصاد القياسي الديناميكية
نماذج الانحدار الذاتي.. ونماذج القيم الموزعة متأخرًا

DYNAMIC ECONOMETRIC MODELS: AUTOREGRESSIVE AND DISTRIBUTED–LAG MODELS

في نماذج الانحدار التي تكون فيها البيانات عبارة عن بيانات سلاسل زمنية، إذا كان نموذج الانحدار يشتمل ليس فقط على القيم الحالية للمتغيرات المفسرة (X's) وإنما أيضًا على القيم المتأخرة (في الماضي) يسمى النموذج بنموذج القيم الموزَّعة متأخرًا، وإذا كان النموذج يحتوي على قيمة أو أكثر متأخرة للمتغير التابع تؤخذ على إنها متغيرات مفسرة، فإن النموذج يطلق عليه نموذج الانحدار الذاتي، فنجد التالي :

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$$

يمثل نموذج القيم الموزعة متأخرًا ، في حين أن :

$$Y_t = \alpha + \beta X_t + \gamma Y_{t-1} + u_t$$

يعتبر مثالاً على نموذج انحدار ذاتي. والأخير يسمى أيضًا النموذج الديناميكي، حيث يصف الشكل الزمني للمتغير التابع، وعلاقته بقيمة أو أكثر من قيم نفس المتغير التابع في الماضي.

نماذج الانحدار الذاتي، ونماذج القيم الموزعة متأخرًا، تستخدم كثيرًا في تحليل الاقتصاد القياسي. وفي هذا الفصل، سندرس بالتفصيل مثل هذه النماذج، ونحاول معرفة التالي :

1 – ما دور القيم المتأخرة في الاقتصاد.

2 – ما هي أسباب وجود القيم المتأخرة.

3 – هل هناك تفسير نظري لنماذج القيم المتأخرة المستخدمة كثيرًا في الاقتصاد القياسي العملي؟

4 – ما هي العلاقة، إن وجدت أي علاقة، بين نماذج الانحدار الذاتي، ونماذج القيم الموزعة متأخرًا؟ هل يمكن اشتقاق واحدة منهما من الأخرى؟

5 – ما هي بعض المشاكل الإحصائية المرتبطة بالتقدير في مثل هذه النماذج؟

6 – هل العلاقة بين قيم المتغيرات الحالية والماضية تمثل علاقة سببية؟ إذا كان ذلك صحيحًا، كيف يمكن قياس ذلك؟

1.17 الدور الذي يلعبه "الزمن" أو "القيم المتأخرة" في الاقتصاد: THE ROLE OF "TIME", OR "LAG" IN ECONOMICS

في الاقتصاد نادرًا ما يكون اعتماد المتغير التابع Y على المتغيرات الأخرى X (المتغيرات المفسرة) في نفس اللحظة الزمنية. بل غالبًا ما تستجيب Y للمتغير X بعد انقضاء فترة زمنية معينة. هذه الفترة الزمنية تسمى فترة متأخرة (Lag). لتوضيح طبيعة هذه الفترة المتأخرة، دعنا نستعرض عدة أمثلة.

مثال 1.17

دالة الاستهلاك : The consumption function

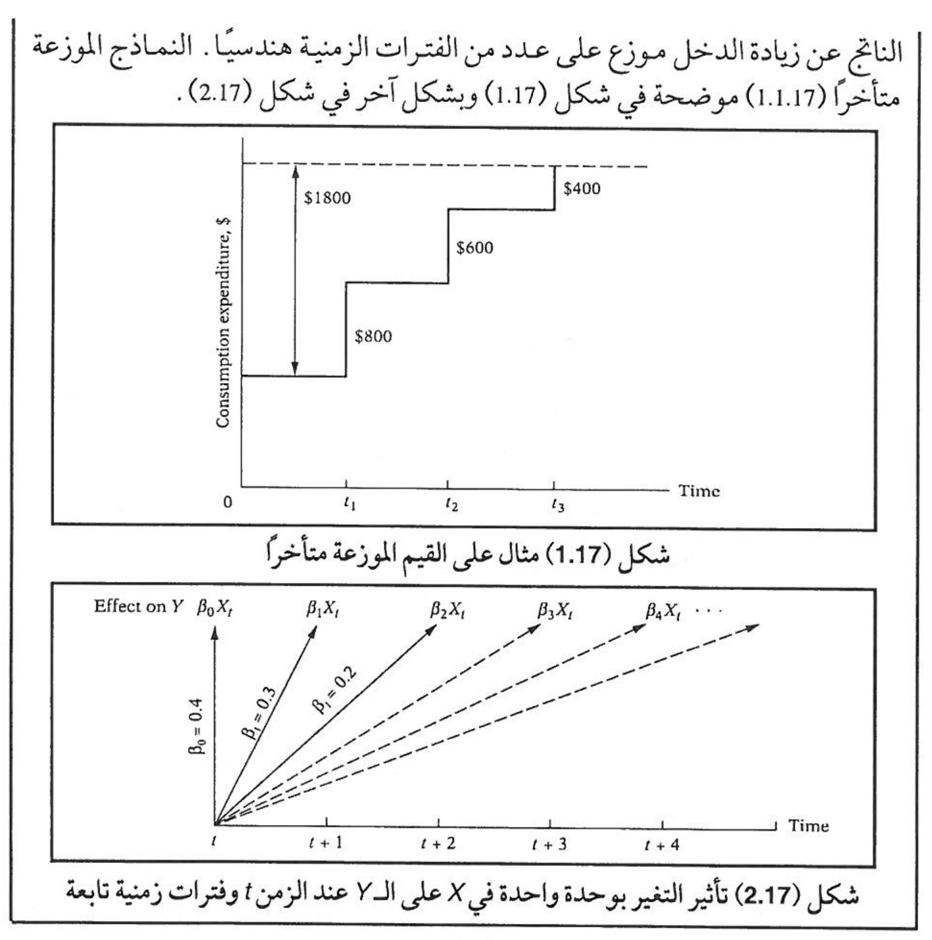
افترض أن هناك شخصًا حصل على زيادة سنوية في الدخل بمقدار 2000$، وافترض أن هذه الزيادة دائمة، بمعنى استمرار حصوله على مثل هذه الزيادة دائمًا. ما أثر هذه الزيادة في الدخل ومصاريف الاستهلاك الخاصة بهذا الفرد سنويًا؟ كنتيجة لهذه الزيادة في الدخل، فإن الأفراد عادة لايسرعون لصرف هذه الزيادة مباشرة. بل إن هذا الفرد قد قرر زيادة مصاريف الاستهلاك بـ 800$ في السنة الأولى كنتيجة لزيادة الدخل، وفي السنة التالية زادت مصاريف الاستهلاك بـ 600$، والسنة التالية بـ 400$، ويدخر المتبقي في نهاية السنة الثالثة، وبالتالي سيزداد استهلاك الفرد السنوي بمقدار 1800$. يمكن التعبير عن دالة الاستهلاك هذه كالتالي:

$$Y_t = \text{constant} + 0.4X_t + 0.3X_{t-1} + 0.2X_{t-2} + u_t \qquad (1.1.17)$$

بحيث إن Y مصاريف الاستهلاك، و X هي الدخل.

المعادلة (1.1.17) تظهر أن أثر زيادة الدخل بـ 2000$ مشتت أو موزع على 3 سنوات، نماذج مثل (1.1.17) تسمى لذلك بالنماذج الموزعة متأخرًا، حيث إن الأثر

الناتج عن زيادة الدخل موزع على عدد من الفترات الزمنية هندسيًا. النماذج الموزعة متأخرًا (1.1.17) موضحة في شكل (1.17) وبشكل آخر في شكل (2.17).



شكل (1.17) مثال على القيم الموزعة متأخراً

شكل (2.17) تأثير التغير بوحدة واحدة في X على الـ Y عند الزمن t وفترات زمنية تابعة

وبشكل عام يمكن كتابة النموذج كالتالي :

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + u_t \qquad \textbf{(2.1.17)}$$

وهذا هو النموذج الموزع متأخرًا بعدد محدود من الفترات المتأخرة زمنيًا k المعامل β_0 معروف باسم الأجل القصير أو التأثير أو المضروب، حيث إنه يعطي التغير في القيمة المتوسطة للـ Y تابعة التغير في X بمقدار الوحدة في نفس الفترة الزمنية(1).

إذا ظل التغير في X على نفس المستوى بعد ذلك، فإن $(\beta_0 + \beta_1)$ يعطي التغير في القيمة المتوقعة للـ Y في الفترة التالية و $(\beta_0 + \beta_1 + \beta_2)$ في الفترة التالية للفترة السابقة

(1) فينا، β_0 هو التفاضل الجزئي لـ Y بالنسبة لـ X_t، β_1 بالنسبة X_{t-1}، β_2 بالنسبة لـ X_{t-2} وهكذا، وبالرموز هذا يعني أن : $\frac{\partial Y_t}{\partial X_{t-k}} = \beta_k$

وهكذا. وهذا التجميع التجزيئي يسمى مرحلي أو متوسط أو مضروب. أخيراً بعد مرور k فترة زمنية نحصل على:

$$\sum_{i=0}^{k}\beta_i = \beta_0 + \beta_1 + \beta_2 + \cdots + \beta_k = \beta \qquad (3.1.17)$$

وهذا معروف باسم المدى البعيد أو المجموع أو مضروب القيم الموزعة متأخراً، وبافتراض وجود هذا المجموع ـ (سنناقش ذلك في مجال آخر)، فإننا نجد التالي:

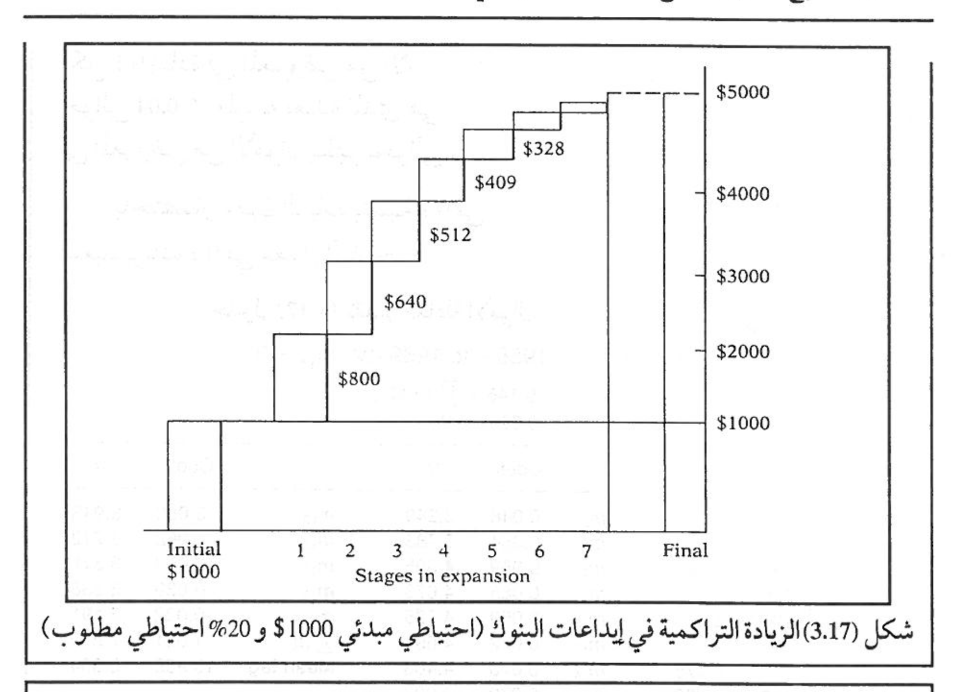
$$\beta_i^* = \frac{\beta_i}{\sum\beta_i} = \frac{\beta_i}{\beta} \qquad (4.1.17)$$

فنعرف β_i على أنه معامل قياسي. والمجاميع الجزيئية للـ β_i القياسي تعطي نسبة المدى البعيد أو المجموع أو التأثير الحادث خلال فترة زمنية ما. بالعودة إلى انحدار الاستهلاك (1.1.17) نجد أن المضروب في المدى القصير والذي ليس إلا الميل الحدى في المدى القصير للاستهلاك (MPC) وهو يساوي 0,4، في حين أن مضروب المدى البعيد والذي هو عبارة عن الميل الحدي في المدى البعيد للاستهلاك يساوي 0,2 + 0,3 + 0,4 = 0,9 وبالتالي فإنه لكل زيادة 1$ في الدخل، فإن المستهلك يزيد من مستوى الاستهلاك بحوالي 40 ¢ في سنة الزيادة و 30¢ أخرى في السنة التالية ثم 20$ أخرى في السنة التابعة للسنة الأخيرة وبالتالي فإن التأثير بعيد المدى لزيادة الدخل بمقدار 1 $ هو 90¢ على الاستهلاك. إذا حصلنا على خارج قسمة β_i على 0,4 نحصل بالترتيب على 0.44 ثم 0.33 و 0.23 مما يعني أن 44% من التأثير الكلي للتغير بمقدار الوحدة في X على Y يحدث مباشرة ثم 77% بعد مرور عام واحد ثم 100% في نهاية السنة الثانية.

مثال 2.17

توليد أموال البنوك (طلب الإيداع) : Creation of bank money (Demand deposits)

افترض أن النظام الاحتياطي الفيدرالي وضع 1000$ من أموال جيدة للنظام البنكي عن طريق شراء بعض الممتلكات الحكومية. ما هي قيمة إجمالي نقود البنوك أو طلب الإيداع الذي سيتولد نتيجة لذلك؟ باتباع نظام الاحتياطي الجزيئي إذا افترضنا أن القانون يطلب من البنوك الاحتفاظ بنسبة 20% كمرجعية للإيداع الموضوع، فبالتالي بعملية المضروب المعروفة، فإن الكمية الكلية لطلب الإيداع التي ستحدث ستكون مساوية لـ $5000 = [1/(1−0.8)] 1000 $. بالطبع فإن 5000$ في طلب الإيداع لن تحدث في يوم وليلة. العملية ستأخذ بعض الوقت، ويمكن توضيح ذلك تخطيطياً في شكل (3.17).



شكل (3.17) الزيادة التراكمية في إيداعات البنوك (احتياطي مبدئي 1000 $ و 20% احتياطي مطلوب)

مثال 3.17

العلاقة بين النقود والأسعار : Link between money and prices

وفقًا للماليين، فإن التضخم هو بالضرورة ظاهرة مالية، حيث إنه عبارة عن زيادة مستمرة في المستوى العام للأسعار والناتج عن زيادة المعروض من الأموال عن كمية الأموال المطلوبة فعليًا من الوحدات الاقتصادية. بالطبع فإن هذه العلاقة بين التضخم والتغير في المعروض من الأموال لايأتي لحظيًا. الدراسات أوضحت أن الفترة الزمنية بين الاثنين هي ما بين 3 إلى 20 فترة ربع سنوية. نتائج إحدى هذه الدراسات موضحة في جدول (1.17)[2]، حيث نرى تأثير التغير بـ 1% في المعروض من الأموال (M1B) (= العملة + إيداع الشيكات في المراكز الاستثمارية المختلفة) في خلال 20 فترة ربع سنوية. التأثير بعيد المدى لتغير المعروض من الأموال لمتغيرات 1% على التضخم حوالي 1 ($=\sum m_i$) وهو معنوي إحصائيًا، في حين أن التأثير قصير المدى بحوالي 0.04 وهو ليس تأثيرًا معنويًا إحصائيًا، على الرغم من أن المضروب الوسيط يبدو كأنه عمومًا معنوي. بالمصادفة نلاحظ أن كلاً من M ، P في شكل نسبة، وبالتالي فإن m_i (المماثل لـ β_i في الترميز المعتاد عليه) تعطي مرونة P بالنسبة لـ M، وهي عبارة عن التغير النسبي للأسعار لكل زيادة 1% في المعروض من الأموال. وبالتالي فإن $m_0 = 0.041$ تعني أن

(2) Keith M.Carlsion, "the lag from money to prices", Review, Federal reserve bank of St. Louis, October, 1980, table 1, p.4.

لكل 1% زيادة في المعروض من الأموال، فإن المرونة السعرية في المدى القصير تكون حوالي 0,04%. المرونة بعيدة المدى هي 1.03% بمعنى أنه على المدى البعيد فإن 1% زيادة في المعروض من الأموال يظهر بحوالي نفس النسبة في الزيادة في الأسعار.

باختصار، فإن الزيادة بنسبة 1% في المعروض من الأموال مصاحب في المدى البعيد بزيادة 1% في معدل التضخم.

جدول (1.17) تقدير معادلة الأموال- الأسعار : تحديد أصلي

Sample period: 1955–I to 1969–IV: $m_{12} = 0$

$$\dot{P} = \underset{(0.395)}{-0.146} + \sum_{i=0}^{20} m_i \dot{M}_{-i}$$

	Coeff.	$\|t\|$		Coeff.	$\|t\|$		Coeff.	$\|t\|$
m_0	0.041	1.276	m_8	0.048	3.249	m_{16}	0.069	3.943
m_1	0.034	1.538	m_9	0.054	3.783	m_{17}	0.062	3.712
m_2	0.030	1.903	m_{10}	0.059	4.305	m_{18}	0.053	3.511
m_3	0.029	2.171	m_{11}	0.065	4.673	m_{19}	0.039	3.338
m_4	0.030	2.235	m_{12}	0.069	4.795	m_{20}	0.022	3.191
m_5	0.033	2.294	m_{13}	0.072	4.694	$\sum m_i$	1.031	7.870
m_6	0.037	2.475	m_{14}	0.073	4.468	**Mean lag**	10.959	5.634
m_7	0.042	2.798	m_{15}	0.072	4.202			
$\bar{R}^2$	0.525							
se	1.066							
D.W.	2.00							

الرموز : P = معدل التغير السنوي في GNP المحسوب

M= معدل التغير السنوي المحسوب في M1B

المصدر: Keith M. Carison, "The lag from money to price", Review, federal reserve bank of st. Louis, October 1980, table 1, P.4.

مثال 4.17

الفترة الزمنية المتأخرة بين مصروفات R&D والإنتاجية :

Lag between R&D expenditure and productivity

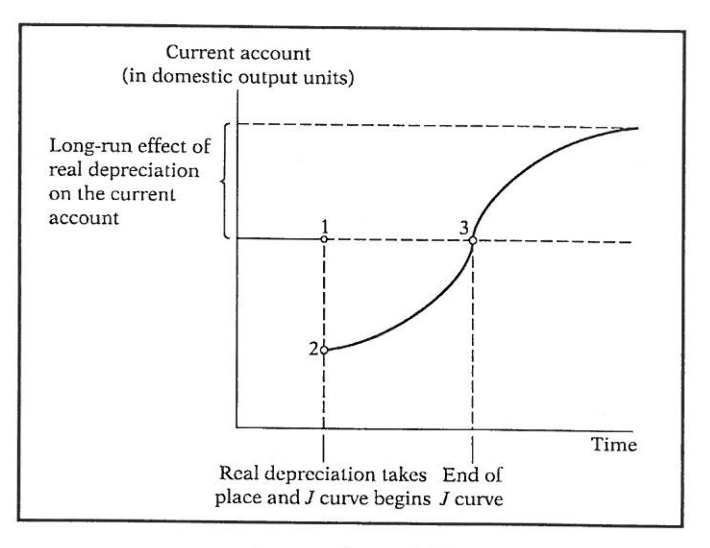
قرار الاستثمار في مصروفات الأبحاث والتطوير (R&D)، ونتيجة ذلك كزيادة في الإنتاج نحتاج لمرور فترة زمنية أو بالأصح عدة فترات زمنية كالتالي، ". . الفترة الزمنية بين الاستثمار في التمويل، ووقت ظهور الاختراعات كنتائج لهذا الاستثمار، الفترة الزمنية بين اختراع فكرة أو جهاز ما ومرحلة ترويجه تجاريًا، والفترة الزمنية لتقديمه أو إدخاله في العملية الإنتاجية، حيث يلزم مرور فترة زمنية قبل استبدال الآلات القديمة بالآلات الجديدة الأفضل"(3).

(3) Zvi Griliches, "Distributed lags, "Econometrica, vol. 36, no.1, January 1967, pp. 16-49.

مثال 5.17

منحنى J للاقتصاد الدولي : The J curve of international economics

طلاب الاقتصاد الدولي على علم بالمقصود بمنحنى J ، والذي يوضح العلاقة بين التوازن التجاري وانخفاض قيمة العملة. فبالنظر إلى انخفاض قيمة العملة في بلد ما (مثل تخفيض قيمة العملة) مبدئيًا، فإن التوازن التجاري يتدهور، ولكن بمرور الوقت يتحسن بافتراض ثبات باقي العوامل الأخرى. المنحنى كما هو موضح في الشكل التالي:



شكل (4.17) منحنى J

المصدر : Paul R.Krugman and Maurice obst feld , international economics: theory and practice, 3d ed., Harper Coilins, New York, 1991, p.465.

مثال 6.17

النموذج التعجيلي للاستثمار : The Accelerator Model of Investment

في أبسط صور النموذج ، فإن مبدأ التعجيل في نظرية الاستثمار يعتمد على أن الاستثمار نسبي التغير في الناتج، بالرموز ذلك يمثل التالي:

$$I_t = \beta(X_t - X_{t-1}) \qquad \beta > 0 \qquad (5.1.17)$$

بحيث إن I_t هو الاستثمار عند الزمن t، X_t هو الناتج عند الزمن t، و X_{t-1} هو الناتج عند الزمن $(t-1)$.

2.17 أسباب الفترات الزمنية المتأخرة(4):

THE REASONS FOR LAGS

على الرغم من أن الأمثلة السابقة في الفقرة 1.17 أوضحت طبيعة ظاهرة الفترات الزمنية المتأخرة، فإنه لم يتم استعراض بشكل كاف أسباب وجود مثل هذه الفترات الزمنية. فهناك ثلاثة أسباب رئيسية وهي :

1- أسباب نفسية : كنتيجة لقوة العادة (الكسل)، فإن الأفراد لايغيرون مباشرة نمط استهلاكهم كنتيجة لانخفاض أو زيادة الأسعار، وقد يكون ذلك بسبب أن عملية تغيير الاستهلاك تشتمل على بعض من عدم المنفعة الوسيطة. فمثلاً الأشخاص الذين يصبحون مليونيرات بعد الفوز في مسابقة اليانصيب، قد لا يغيرون نمط حياتهم مباشرة، حيث اعتادوا على حياة مختلفة تماماً لفترات زمنية طويلة، وقد يكون من الصعب عليهم عمل رد فعل سريع لوجود مثل هذه الثروة الهابطة فجأة من السماء.

بالطبع مع مرور وقت كاف قد يتعلمون كيفية التعامل مع هذا الموقف، وبدء حياة جديدة متماشية مع مثلٍ هذه الثروة. أيضاً قد لا يعلم الأفراد ما إذا كان هذا التغير دائماً أم مؤقتاً. وبالتالي فإن رد الفعل على الزيادة في الدخل يعتمد على كون هذه الزيادة مستمرة أم مؤقتة. فإذا كانت هذه الزيادة مؤقتة ومقصورة على فترة نجاح ما، وسيعود الدخل إلى ما كان عليه في الماضي، فإنه من المنطقي لشخص ما أن يدخر الزيادة في الدخل، وقد يرى البعض الآخر أنه من المنطقي أكثر صرف كل هذه الزيادة المؤقتة.

2- أسباب فنية : بافتراض أن سعر رأس المال بالنسبة للعمالة ينخفض، مما يجعل تبديل رأس المال بالعمالة ممكناً اقتصادياً. بالطبع إضافة المزيد من رأس المال يحتاج إلى وقت (مرحلة التكوين)، بالإضافة إلى ذلك إذا كان انخفاض السعر مؤقتاً فإن المؤسسة قد لاتسرع في تبديل رأس المال بالعمالة، خصوصاً إذا كان من المتوقع بعد هذا الانخفاض المؤقت في سعر رأس المال أن يعود ويزداد أكثر من مستوياته السابقة. فأحياناً عدم توافر معلومات كافية يكون أحد أسباب وجود فترات زمنية متأخرة lags، في الوقت الحاضر سوق الحاضر سوق الحاسب الإليكتروني الشخصي ممتلئة بالعديد من الأنواع المختلفة للحاسبات الإلكترونية التي لها صفات مختلفة وأسعار متباينة. بالإضافة إلى ذلك، عندما كانت البداية في أواخر

(4) This section leans heavily on Marc Nerlove, Distributed Lags and Demand Analysis for Agricultural and Other Commoditites, Agricultural Handbook No. 141, U.S. Department of Agriculture, June 1958.

السبعينيات 1970، فإن أسعار معظم الحاسبات الإلكترونية الشخصية انخفضت بشكل ملحوظ. وكنتيجة لذلك، فإن المستهلكين للحاسبات الشخصية قد يترددون في الشراء حتى يأخذوا الوقت الكافي للبحث في مميزات وأسعار الأنواع العديدة المتنافسة من الحسابات الشخصية، إلى جانب أنهم قد يزايدون في الشراء، متوقعين انخفاض السعر أكثر في المستقبل، وظهور أنواع أكثر حداثة من المتاح الآن.

3- **أسباب مؤسسية**: هذه الأسباب تعد أيضًا من أسباب الفترات الزمنية المتأخرة lags. فعلى سبيل المثال، التزامات التشييد قد تعوق المؤسسة من تغير العمالة أو المواد الخام. وكمثال آخر، الأشخاص الذين يستثمرون أموالهم في البنوك كودائع لفترات زمنية طويلة 1 سنة، 3 سنوات أو 7 سنوات يكون من غير الممكن مشاركتهم في أي سوق تجارية حتى ولو كانت مكاسبهم أكثر مما يحصلون عليه من البنك. وبنفس الشكل، فإن العاملين في شركة ما يكون مسموحًا لهم اختيار خطة التأمين الصحي التي يرغبون فيها، ولكن عندما يختار العامل خطة ما يكون غير مسموح له الخروج منها والدخول في خطة أخرى، قبل أن ينقضي على الأقل عام واحد على خطة التأمين الصحي المختارة. قد يكون ذلك جيدًا ولمصلحة إدارة المؤسسة، ولكن بالنسبة للعامل فهو لايستطيع فعل أي تغير قبل مرور عام كامل.

ووفقًا لجميع الأسباب المذكورة سابقًا، فإن الفترات الزمنية المتأخرة تلعب دورًا مهمًا في الاقتصاد. وهذا بالطبع يؤثر على خطط المدى القصير والطويل الاقتصادية. ولهذا السبب، فإننا نقول إن الأسعار في المدى القصير أو مرونة الاقتصاد في المدى القصير عمومًا أقل (القيمة المطلقة) من نظيرها في المدى البعيد. أو بمعنى آخر، فإن الميل الحدى للاستهلاك في المدى القصير عمومًا أقل من الميل الحدي للاستهلاك في المدى البعيد.

3.17 تقدير النماذج الموزعة متأخراً:
ESTIMATION OF DISTRIBUTED-LAG MODELS

مما لاشك فيه، فإن النماذج الموزعة متأخرًا تلعب دورًا مهمًا في الاقتصاد، كيف يمكن تقدير مثل هذه النماذج؟ بالتحديد دعنا نفترض وجود النموذج التالي، والذي يشتمل على متغير مفسر واحد[5].

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + u_1 \qquad (1.3.17)$$

(5) إذا لدينا أكثر من متغير مفسر واحد، وكل متغير قد يكون له تأثير على X في فترات زمنية مختلفة. للتبسيط نفترض وجود متغير مفسر واحد فقط.

ونلاحظ أننا لم نحدد عدد الفترات الزمنية المتأخرة المستخدمة، بمعنى إلى أي حد سنعود إلى قيم المتغير في الماضي. مثل هذا النموذج يسمى نموذج قيم زمنية متأخرة لانهائي، في حين أن نموذج مثل المعطى في (1.2.17) يسمى النموذج الموزع متأخراً المحدود، حيث إن طول الفترات الزمنية k محددة. سنستخدم نموذج (1.3.17) حيث إنه أسهل من الناحية الرياضية كما سنرى لاحقًا[6].

كيف يمكن تقدير قيم α و β's الموجودة في (1.3.17)؟ من الممكن استخدام إحدى الطريقتين التاليتين: (1) تقدير ad hoc و (2) قيود مسبقة على الـ β's بافتراض أن الـ β's تتبع بعض التوزيعات المتماثلة. سنستعرض تقدير ad hoc في الفقرة التالية، والأسلوب الآخر في فقرة 4.17.

تقدير Ad Hoc للنماذج الموزعة متأخراً:

Ad Hoc Estimation of distributed- lag models

بما أن المتغير المفسر X_t يفترض أنه غير عشوائي (non stochastic) (أو على الأقل غير مرتبط مع مقدار الخطأ u_t)، X_{t-1}، X_{t-2} وهكذا مفترض أيضًا أنها مقادير غير عشوائية. وبالتالي مبدئيًا، فإن طريقة المربعات الصغرى العادية (OLS) من الممكن تطبيقها على (1.3.17). وهذا الأسلوب قام به Alt[7] و Tinbergen[8].

هما يقترحان لتقدير (1.3.17) استخدام أسلوب تدريجي، بمعنى أن يقوم أولاً انحدار Y_t على X_t ثم انحدار Y_t على كل من X_t و X_{t-1} ثم انحدار Y_t على كل من X_t، X_{t-1}، X_{t-2} وهكذا. هذه العملية التدريجية تتوقف عندما تصبح معاملات انحدار قيم X في الفترات الزمنية المتأخرة غير معنوية إحصائيًا، أو معامل انحدار واحد على الأقل من المتغيرات تغيرت إشارته من موجب إلى سالب أو العكس. باتباع هذه الطريقة، فإن Alt قام بانحدار استهلاك البترول Y على الطلبات الجديدة X. معتمداً على قيم البيانات الربع سنوية للفترة من 1930–1939 فكانت النتائج كالتالي:

---

(6) في الواقع، عمومًا معاملات قيم الـ X المختلفة من المتوقع أن يكون لها تأثير غير مهم على الـ Y.

(7)F.F.Alt, "Distributed lags", Econometrica, vol. 10, 1942, pp.113-128.

(8) J.Tinbergen, "Long-Term foreign trade Elasticities", Metroeconomica, vol. I, 1949, pp.174-185.

$$\hat{Y}_t = 8.37 + 0.171X_t$$
$$\hat{Y}_t = 8.27 + 0.111X_t + 0.064X_{t-1}$$
$$\hat{Y}_t = 8.27 + 0.109X_t + 0.071X_{t-1} - 0.055X_{t-2}$$
$$\hat{Y}_t = 8.32 + 0.108X_t + 0.063X_{t-1} + 0.022X_{t-2} - 0.020X_{t-3}$$

اختار Alt الانحدار الثاني كأفضل انحدار، حيث إنه في المعادلتين الأخريين إشارة المتغير X_{t-2} لم تكن ثابتة، وفي المعادلة الأخيرة كانت إشارة X_{t-3} والذي يصعب تفسيره اقتصاديًا.

وعلى الرغم من أن تقدير Ad hoc يعتبر مباشرًا نوعًا ما، فإنه يعاني من العديد من المشاكل كالتالي:

1 - توجد مرجعية ثابتة لأقصى عدد ممكن استخدامه من الفترات الزمنية المتأخرة(9).

2 - كلما تم تقدير عدد أكبر من الفترات الزمنية المتأخرة كلما قل عدد درجات الحرية مما يجعل هناك صعوبة في عملية الاستدلال الإحصائي. وعمومًا فالاقتصاديون عادة لايحالفهم الحظ في إيجاد بيانات لفترات زمنية طويلة تجعلهم قادرين على تقدير فترات زمنية متأخرة عديدة.

3 - المشكلة الأكثر أهمية والموجودة بكثرة في بيانات السلاسل الزمنية في الاقتصاد، هي أن قيم المتغيرات المفسرة على فترات زمنية متأخرة، عادة ما تكون مرتبطة مع بعضها البعض، وبالتالي تظهر مشكلة الارتباط الخطي المتعدد. وكما سبق وذكرنا في الفصل (10)، فإن مشكلة الارتباط الخطي المتعدد تجعل التقديرات غير دقيقة، حيث إن الأخطاء القياسية تكون كبيرة بالنسبة للمعادلات المقدرة.

وكنتيجة لذلك، وبالاعتماد على قيمة t المحسوبة، سنميل دائمًا إلى قبول الفرض العدمي، وتكون معاملات الانحدار لقيم المتغير المفسرة في الفترات الزمنية المتأخرة غير معنوية إحصائيًا.

4 - البحث عن طول الفترات الزمنية المتأخرة يفتح المجال أمام الباحثين للتطرق إلى موضوع تنقيب البيانات. أيضًا كما رأينا من قبل في فقرة 4.13 القيمة الإسمية والقيمة الفعلية لمستوى المعنوية المختارة في اختبارات الفروض الإحصائية تلعبان دورًا كبيرًا في هذه الأبحاث التسلسلية.

(9) إذا كان طول الفترة الزمنية المتأخرة k محددة بطريقة خاطئة، سنحتاج إلى التعامل مع مشكلة خطأ التوصيف التي تناولناها في الفصل (13). وضع في اعتبارك التحذير الخاص بالـ data mining

وبالنظر إلى مثل هذه المشاكل، فإن طريقة التقدير Ad hoc لايوصى باستخدامها كثيراً. وبشكل واضح، فإن هناك بعض الاعتبارات النظرية التي يجب الرجوع إليها للاعتماد على قيم β's المقدرة إذا أردنا التغلب على بعض مشاكل التقدير بهذه الطريقة.

4.17 أسلوب KOYCK للنماذج الموزعة متأخراً:
THE KOYCK APPROACH TO DISTRIBUTED-LAG MODELS

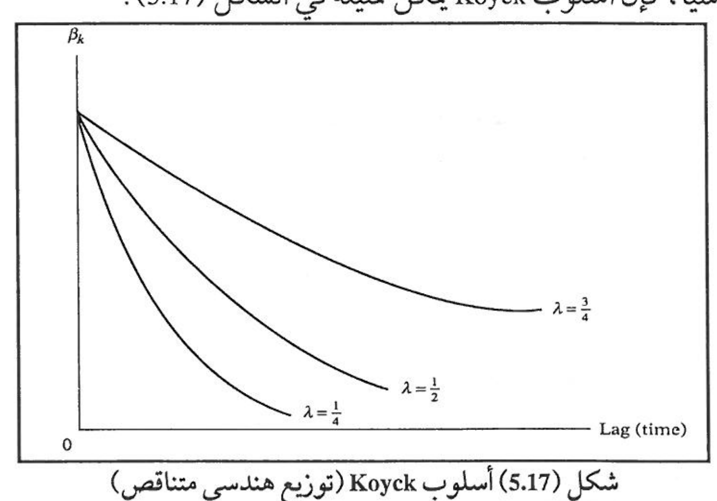
اقتراح Koyck طريقة ماهرة لتقدير النموذج الموزع متأخراً. دعنا نبدأ بالنموذج الموزع متأخراً اللانهائي (1.3.17). بافتراض أن الـ β's جميعاً لها نفس الإشارة Koyck افترض تناقصها هندسياً كالتالي[10].

$$\beta_k = \beta_0 \lambda^k \qquad k = 0, 1, \ldots \qquad [11](1.4.17)$$

بحيث إن λ هي معدل التناقص، وتكون $0 < \lambda < 1$ ويطلق على $1-\lambda$ سرعة التسوية.

المقصود من (1.4.17) أن كل معامل β يكون أقل عددياً من معامل β السابق له (يتحقق ذلك عندما تكون $\lambda<1$) بما يعني أنه بالعودة إلى الوراء في الزمن، تأثير الفترات الزمنية المتأخرة على Y_i يصبح أقل، وهذا الفرض مقبول إلى حد كبير عملياً. فعموماً الدخل الحالي والدخل في السنوات الماضية القريبة من المتوقع أن يكون له تأثير على الاستهلاك الحالي عن تأثير الدخل في السنوات الماضية البعيدة على الاستهلاك الحالي.

هندسياً، فإن أسلوب Koyck يمكن تمثيله في الشكل (5.17).



شكل (5.17) أسلوب Koyck (توزيع هندسي متناقص)

(10) L.M.Koyck, "Distributed lags and investment analysis, North Holland publishint company, Amsterdam, 1954.

(11) أحياناً تتم كتابة هذه المعادلة كالتالي:

$$\beta_k = \beta_0(1-\lambda)\lambda^k \qquad k = 0, 1, \ldots$$

والأسباب معطاة في الملحوظة 12.

كما هو موضح في الشكل السابق، قيمة معامل β_k للفترات الزمنية المتأخرة تعتمد على القيمة λ بغض النظر عن القيمة المشتركة β_0. فكلما اقتربت λ من 1 كلما قل معدل انخفاض β_k في حين كلما اقتربت λ من الصفر، كلما زادت سرعة الانخفاض في β_k. في الحالة المثالية، فإن القيم الماضية القريبة للـ X سيكون لها تأثير ملحوظ على Y_t، في حين قيم X الماضية البعيدة سيتلاشى أثرها على الـ Y سريعًا. وهذا النمط يتضح كالتالي:

λ	β_0	β_1	β_2	β_3	β_4	β_5	···	β_{10}
0.75	β_0	$0.75\beta_0$	$0.56\beta_0$	$0.42\beta_0$	$0.32\beta_0$	$0.24\beta_0$	···	$0.06\beta_0$
0.25	β_0	$0.25\beta_0$	$0.06\beta_0$	$0.02\beta_0$	$0.004\beta_0$	$0.001\beta_0$	···	0.0

لاحظ الصفات التالية المميزة لأسلوب Koyck: (1) بافتراض القيم غير السالبة للـ λ، Koyck يفترض عدم تغير إشارة الـ β's و (2) بافتراض λ<1 فإنه يعطي أوزانًا أقل للـ β's المرتبطة بفترات زمنية بعيدة عن الـ β's المرتبطة بالقيم الحالية و (3) يفترض أن مجموع الـ β's والذي يمثل المضروب في المدى البعيد محدود ويساوي:

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \left(\frac{1}{1-\lambda}\right) \qquad \text{(12)}(2.4.17)$$

وكنتيجة لـ (1.4.17) يكون نموذج الفترات الزمنية المتأخرة اللانهائي (1.3.17) يمكن كتابته كالتالي:

$$Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \cdots + u_t \qquad (3.4.17)$$

وكما هو واضح، فإن النموذج مازال غير قابل للتعديل، بحيث يسمح بالتقدير السهل، حيث إن هناك عددًا كبيرًا من المعاملات (اللانهائية) يجب تقديره، والمعامل λ في شكل غير خطي عالي التعقيد: وبشكل محدد، فإن أسلوب تحليل الانحدار الخطي (الخطي في المعالم) لايمكن تطبيقه في مثل هذه الحالة.

ولكن Koyck يقترح الآن طريقة للتغلب على ذلك. حيث يقترح استخدام فترة زمنية واحدة متأخرة كما في (3.4.17) للحصول على:

$$Y_{t-1} = \alpha + \beta_0 X_{t-1} + \beta_0 \lambda X_{t-2} + \beta_0 \lambda^2 X_{t-3} + \cdots + u_{t-1} \qquad (4.4.17)$$

---

(12) هذا بسبب

$$\sum \beta_k = \beta_0(1 + \lambda + \lambda^2 + \lambda^3 + \cdots) = \beta_0 \left(\frac{1}{1-\lambda}\right)$$

وبما أن المقدار الموجود بين الأقواس هو متسلسلة هندسية لانهائية، فإن مجموعها هو $\left(\frac{1}{1-\lambda}\right)$ و 0<λ<1. كما سبق وذكرنا، إذا كانت β_k معرفة كما هو موجود في الملاحظة 11 فإن $\sum\beta_k = \beta_0(1-\lambda)/(1-\lambda) = \beta_0$ مما يضمن أن الأوزان $(1-\lambda)\lambda^k$ تتجمع إلى واحد صحيح.

ثم بضرب المعادلة (4.4.17) في λ فيحصل على :

$$\lambda Y_{t-1} = \lambda\alpha + \lambda\beta_0 X_{t-1} + \beta_0\lambda^2 X_{t-2} + \beta_0\lambda^3 X_{t-3} + \qquad (5.4.17)$$

وبطرح (5.4.17) من (3.4.17) يحصل على :

$$Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + (u_t - \lambda u_{t-1}) \qquad (6.4.17)$$

وبإعادة الترتيب نحصل على :

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t \qquad (7.4.17)$$

حيث :

$v_t = u_t - \lambda u_{t-1}$ ويسمى متوسط متحرك لكل من u_t و u_{t-1} .

الطريقة التي ذكرت الآن تسمى تحويلة Koyck. بمقارنة (7.4.17) مع (1.3.17) سنرى التبسيط الهائل الذي قام به Koyck. حيث كان من قبل هناك ضرورة لتقدير α وعدد لانهائي من الـ β's، أما الآن نحتاج على تقدير ثلاثة مجاهيل فقط : α، β_0 و λ. والآن لايوجد أي سبب لتوقع وجود ارتباط خطي متعدد. بمعنى أن مشكلة الارتباط الخطي المتعدد تم التغلب عليها بتبديل .. $X_{t-1}, X_{t-2},$ بقيمة واحدة فقط وهي Y_{t-1}. ولكن لاحظ الصفات التالية لتحويلة Koyck:

1 - بدأنا بالنموذج الموزع متأخراً، ولكن انتهينا بنموذج انحدار ذاتي، حيث Y_{t-1} تظهر كأحد المتغيرات المفسرة. هذه التحويلة توضح كيف يمكن تقريب أو تحويل النموذج الموزع متأخراً إلى نموذج انحدار ذاتي.

2 - ظهور Y_{t-1} قد يؤدي إلى بعض المشاكل الإحصائية Y_{t-1} مثل Y_t متغير عشوائي، بمعنى أن لدينا متغيراً مفسراً في النموذج عبارة عن متغير عشوائي. نتذكر أن النظرية التقليدية للمربعات الصغرى تقوم على أساس افتراض أن المتغيرات المفسرة إما غير عشوائية، وإن كانت عشوائية، فيجب أن تكون مستقلة عن مقدار الخطأ العشوائي. وبالتالي يجب أن نرى ما إذا كانت Y_{t-1} محققة لهذا الشرط أما لا. (سنعود إلى هذه النقطة في الفقرة 8.17).

3 - في النموذج (1.3.17) مقدار التشتيت (الخطأ) هو v_t في حين أنه في النموذج المحول هو $v_t = (ut - \lambda u_{t-1})$. الخصائص الإحصائية لـ v_t تعتمد على ما هو مفترض من خصائص لـ u_t وبالتالي، كما سيوضح لاحقًا، إذا كان المقدار u_t's أصلاً يوجد فيه عدم ارتباط تسلسلي، فإن u_t's هو أيضًا غير مرتبط تسلسليًا. وبالتالي يجب

النظر إلى مشكلة الارتباط التسلسلي بالإضافة إلى المتغير المفسر العشوائي Y_{t-1}. وسنقوم بدراسة ذلك في الفقرة 8.17.

4 - وجود قيم في فترات زمنية متأخرة للـ Y يخالف أحد الشروط الخاصة بإحصاء Durbin- Watson (d). وبالتالي يجب البحث عن بديل لاختيار الارتباط التسلسلي في حالة وجود قيمة متأخرة لـ Y، أحد هذه البدائل هو اختيار (h) Durbin والذي سيناقش في الفقرة 10.17.

كما رأينا في (4.1.17) التجميع الجزيئي لقيمة β_i القياسية يخبرنا بنسبة التأثير الكلي أو التأثير في المدى البعيد الحادث في فترة زمنية معينة. في الواقع يتم استخدام متوسط أو وسيط الفترات الزمنية المتأخرة كتعبير عن الطبيعية التكوينية لهذه الفترات الزمنية المتأخرة والموجودة في النموذج الموزع متأخراً.

وسيط الفترات الزمنية المتأخرة: The Median Lag

وسيط الفترات الزمنية المتأخرة، هو الزمن المطلوب لحدوث نصف أو 50% من التغير الكلي في Y والتابع لتغير X بوحدة واحدة. في نموذج Koyck، وسيط الفترات الزمنية المتأخرة يتبع التالي (انظر تمرين 6.17):

$$\text{Koyck model: Median lag} = -\frac{\log 2}{\log \lambda} \qquad (8.4.17)$$

وبالتالي، إذا كانت $\lambda = 0.2$، فإن وسيط الفترات الزمنية المتأخرة هو 0.4306، ولكن إذا كانت $\lambda = 0.8$ فإن وسيط الفترات الزمنية المتأخرة يكون 3.1067، بمعنى آخر، فإن 50% من التغير الكلي في Y يحدث قبل مرور نصف المدة، في حين في الحالة الأخرى، فإننا نحتاج إلى 3 فترات حتى نصل إلى 50% من التغير، وهذه النتيجة غير مفاجئة، حيث نعلم أنه كلما زادت قيمة λ قلت سرعة التعديل، وكلما قلت قيمة λ كلما زادت سرعة التعديل.

متوسط الفترات الزمنية المتأخرة: The Mean Lag

باعتبار β_k جميعاً قيماً موجبة، فإن القيمة المتوقعة أو المتوسط الخاص بالفترات الزمنية المتأخرة يعرف كالتالي:

$$\text{Mean lag} = \frac{\sum_0^\infty k\beta_k}{\sum_0^\infty \beta_k} \qquad (9.4.17)$$

وهو ببساطة المتوسط المرجح لكل الفترات الزمنية الموجودة في النموذج مع معاملات الانحدار الخاصة بكل فترة، والموجودة في صورة أوزان. باختصار يعتبر ذلك متوسطًا مرجحًا للفترات الزمنية المتأخرة. في نموذج Koyck متوسط الفترات الزمنية المتأخرة (انظر تمرين 7.17) هو:

$$\text{Koyck model: Mean lag} = \frac{\lambda}{1-\lambda} \qquad (10.4.17)$$

وبالتالي، إذا كانت $\lambda = \frac{1}{2}$ فإن متوسط الفترات الزمنية المتأخرة يساوي 1.

مما سبق يتضح أن وسيط ومتوسط الفترات الزمنية المتأخرة ممكن اعتباره مقياسًا لسرعة استجابة Y للتغير في X. في المثال المعطى في جدول (1.17) متوسط الفترات الزمنية المتأخرة هو 11 فترة ربع سنوية، مما يوضح أننا نحتاج إلى مرور بعض الوقت، في المتوسط، حتى يبدأ التغيير في المعروض من الأموال في التأثير في تغييرات الأسعار.

مثال 7.17

معدل الاستهلاك الشخصي : Per capital personal consumption

في هذا المثال، تتم دراسة الاستهلاك الشخصي لكل فرد (PPCE) وعلاقته بالدخل الشخصي (PPDI) في الولايات المتحدة الأمريكية في الفترة 1970–1999. كل البيانات مقيدة بالدولارات في 1996. لتوضيح نموذج Koyck، اعتبر البيانات المعطاة في جدول (2.17) انحدار PPCE على PPDI والفترات الزمنية المتأخرة لـ PPCE يعطي النتائج التالية:

$$\begin{aligned}\widehat{PPCE}_t &= -1242.169 + 0.6033PPDI_t + 0.4106PPCE_{t-1}\\ se &= (402.5784) \quad (0.1502) \quad (0.1546)\\ t &= (-3.0855) \quad (4.0155) \quad (2.6561)\\ R^2 &= 0.9926 \quad d = 1.0056 \quad \text{Durbin } h = 5.119\end{aligned} \qquad (11.4.17)$$

لاحظ أن: حساب Durbin h تمت مناقشته من قبل في فقرة 10.17 إذا افترضنا أن هذا النموذج يأتي كنتيجة لتحويله من النوع Koyck و λ تساوي 0.4106 وبالتالي فإن وسيط الفترات الزمنية المتأخرة هو:

$$-\frac{\log(2)}{\log\lambda} = -\frac{\log(2)}{\log(0.4106)} = 0.7786$$

ومتوسط الفترات الزمنية المتأخرة هو :

$$\frac{\lambda}{1-\lambda} = \frac{0.4106}{0.5894} = 0.6966$$

بمعنى آخر، يتضح أن PPCE يتأثر بـ PPDI في فترة زمنية قصيرة نسبيًا.

جدول (2.17)

Observation	PPCE	PPDI	Observation	PPCE	PPDI
1970	11,300	12,823	1985	16,020	18,229
1971	11,581	13,218	1986	16,541	18,641
1972	12,149	13,692	1987	16,398	18,870
1973	12,626	14,496	1988	17,463	19,522
1974	12,407	14,268	1989	17,760	19,833
1975	12,551	14,393	1990	17,899	20,058
1976	13,155	14,873	1991	17,677	19,919
1977	13,583	15,256	1992	17,989	20,318
1978	14,035	15,845	1993	18,399	20,384
1979	14,230	16,120	1994	18,910	20,709
1980	14,021	16,063	1995	19,294	21,055
1981	14,069	16,265	1996	19,727	21,385
1982	14,105	16,328	1997	20,232	21,838
1983	14,741	16,673	1998	20,989	22,672
1984	15,401	17,799	1999	21,901	23,191

لاحظ أن: PPCE= نفقات الاستهلاك الشخصي لكل فرد، مقيد بالدولارات 1996

PPDI= الدخل الشخصي لكل فرد، مقيد بدولارات 1996

المصدر: Economic Report of the president, 2001, Table B-31, p.311

5.17 نموذج KOYCK الرشيد.. نموذج التوقعات المتكيفة :

RATIONALIZATION OF THE KOYCK MODEL: THE ADAPTIVE EXPECTATION MODEL

على الرغم من تميز نموذج Koyck (7.4.17) فهو نموذج من نوع ad hoc، حيث إنه تم الحصول عليه من طريقة جبرية ضعيفة، فهو مجرد من أي خلفية نظرية، ولكن هذه الفجوة يمكن التغلب عليها إذا بدأنا النظر للموضوع بطريقة مختلفة. إذا افترضنا النموذج التالي:

$$Y_t = \beta_0 + \beta_1 X_t^* + u_t \tag{1.5.17}$$

بحيث إن Y = الطلب على الأموال (مقدار السيولة)

X^* = تغطية التوازن، الأمثلية، معدل الفائدة المتوقع للمدى البعيد

u = مقدار الخطأ

معادلة (1.5.17) نفترض أن الطلب على المال دالة في معدل الفائدة المتوقع.

بما أن القيمة المتوقعة X^* لا تشاهد مباشرة، دعنا نفترض الفروض التالية عن كيفية تكوين التوقعات:

$$X_t^* - X_{t-1}^* = \gamma(X_t - X_{t-1}^*) \qquad {}^{(13)}(2.5.17)$$

بحيث إن γ، $0 < \gamma \leq 1$، تعرف بأنها معامل التوقع. الفرض (2.5.17) يعرف باسم التوقع المتكيف، التوقع المتدرج أو فرض الخطأ التعليمي. هذه التعريفات مقدمة من كل من Cagan[14] و Friedman[15].

ما تعنيه (2.5.17) هو أن "المؤسسات الاقتصادية ستكيف توقعاتها في ضوء خبراتها الماضية، وبالتحديد ستتعلم من أخطائها"[16]. بتحديد أكثر، فإن (2.5.17) تنص على أن التوقعات في كل فترة بنسبة γ من الفجوة بين القيمة الحالية للمتغير وقيمته المتوقعة السابقة. وبالتالي في نموذجنا الحالي يعني ذلك أن التوقعات عن معدلات الفائدة تنقح في كل فترة بمقدار γ من الفرق بين معدلات الفائدة الموجودة في القيمة الحالية وما هو متوقع في الفترات السابقة كطريقة أخرى للتعبير عن النموذج (2.5.17) كالتالي:

$$X_t^* = \gamma X_t + (1-\gamma)X_{t-1}^* \qquad (3.5.17)$$

والذي يوضح أن القيمة المتوقعة لمعدل الفائدة عند الزمن t هو متوسط مرجح للقيمة الفعلية لمعدل الفائدة عند الزمن t، والقيمة المتوقعة في الفترة السابقة بالأوزان γ و $1-\gamma$ بالترتيب. إذا كانت $\gamma=1$، $X_t^* = X_t$ فإن هذا يعني أن التوقعات يتم إدراكها بالكامل ومباشرة، بمعنى أن التأثير يحدث في نفس الفترة الزمنية.

على العكس إذا كان $\gamma=0$، $X_t^* = X_{t-1}^*$ فإن ذلك يعني أن التوقعات ثابتة، بمعنى أن "الظروف الموجودة في الحاضر ستظل في الفترات المتتابعة التالية، وبالتالي فالقيم المتوقعة المستقبلية ستحدد بالقيم الحالية"[17].

بالتعويض عن (3.5.17) في (1.5.17) نحصل على:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1[\gamma X_t + (1-\gamma)X_{t-1}^*] + u_t \\ &= \beta_0 + \beta_1\gamma X_t + \beta_1(1-\gamma)X_{t-1}^* + u_t \end{aligned} \qquad (4.5.17)$$

(13) أحيانًا يتم التعبير عن النموذج كالتالي:

$$X_t^* - X_{t-1}^* = \gamma(X_{t-1} - X_{t-1}^*)$$

(14)P. Cagan, "The Monetary Dynamics of Hyperinflations," in M. Friedman (ed.), Studies in the Quantity Teory of Money, University of Chicago Press, Chicago, 1956.

(15) Milton Friedman, A Teory of te Consumption Function, National Bureau of Economic Research, Princeton University Press, Princeton, N. J., 1957.

(16) G. K. Shaw, Rational Expectations: An Elementary Exposition, St. Martin's Press, New York, 1984, p. 25.

(17) Ibid., pp. 19-20.

والآن بالعودة لفترة زمنية واحدة متأخرة في (1.5.17) والضرب في $\gamma - 1$ ثم طرح حاصل الضرب من (4.5.17) وبعملية جبرية بسيطة نحصل على:

$$\begin{aligned} Y_t &= \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma)Y_{t-1} + u_t - (1-\gamma)u_{t-1} \\ &= \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma)Y_{t-1} + v_t \end{aligned} \qquad (5.5.17)$$

بحيث أن : $v_t = u_t - (1-\gamma)u_{t-1}$

وقبل الانتقال إلى نقطة أخرى، دعنا نلاحظ الفرق بين (1.5.17) و (5.5.17) في النموذج التقليدي، β_1 تقيس متوسط استجابة للمتغير Y للتغير بمقدار الوحدة في X^*، قيمة X في المدى البعيد أو القيمة التوازنية.

على الجانب الآخر، في (5.5.17) فإن $\gamma\beta_1$ تقيس متوسط استجابة Y لكل وحدة تغير في القيمة الفعلية أو المشاهدة لـ X. هذه الاستجابات لن تكون متساوية إلا إذا كانت بالضرورة $\gamma=1$، بمعنى أن القيم الحالية و قيم X في المدى البعيد متساوية. في الواقع نبدأ بتقدير (5.5.17) وبمجرد حصولنا على تقدير γ من معامل القيم الزمنية المتأخرة للـ Y نستطيع بسهولة حساب β_1 عن طريق قسمة معامل X_t $(=\gamma\beta_1)$ على γ.

التماثل بين نموذج التوقعات المتكيفة (5.5.17) ونموذج Koyck (7.4.17) يجب أن تكون واضحة، حتى ولو كان تفسير المعاملات في النموذجين مختلفًا. لاحظ أنه كما في نموذج Koyck، نموذج التوقعات المتكيفة هو نموذج انحدار ذاتي، ومقدار الخطأ مماثل مع مقدار الخطأ في نموذج Koyck سوف نعود إلى تقدير نموذج التوقعات المتكيفة في فقرة 8.17 وبعض الأمثلة في الفقرة 12.17. والآن بعد أن استعرضنا نموذج التوقعات المتكيفة (AE) إلى أي مدى يعتبر هذا النموذج نموذجًا واقعيًا؟ من الواضح أنه أسلوب أفضل من أسلوب Koyck الضعيف رياضيًا، ولكن هل فرضية AE منطقية؟ بالنسبة لفرض AE من الممكن أن نقول التالي:

إنه يقدم معنى بسيطًا نسبيًا لنمذجة التوقعات في النظرية الاقتصادية، في حين أن فرضية النموذج عن السلوك الخاص بالمؤسسات الاقتصادية يبدو على نحو بارز فرضًا رشيدًا. الاعتقاد بأن الأشخاص يتعلمون من خبراتهم السابقة يعتبر فرضًا عاقلاً عن افتراض أن الأشخاص مجردون من الذاكرة، ويخضعون لفكرة التوقعات الثابتة. بالإضافة إلى أن الفرض الخاص بأن الأحداث في الماضي البعيد لها تأثير أقل من الأحداث في الماضي القريب يعتبر فرضًا مقبولاً، ويمكن التأكد منه من الملاحظة البسيطة اليومية[18].

(18) Ibid., p.27.

قبل ظهور فرض التوقعات الرشيدة (RE) والذي يعود إلى J.Muth ولاحقًا إلى Robert Lucas and Thomas Sargent، كان فرض AE واسع الشهرة في الاقتصاد التجريبي. أساس فرض RE يعتبر امتدادًا لفرض AE حيث يرى فرض RE أن فرض AE غير كامل، حيث إنه يعتمد فقط على القيم الماضية للمتغير لتكوين التوقعات[19]. في حين أن فرض RE يعتمد على أن مفردة المؤسسة الاقتصادية تعتمد على المعلومات المتاحة حاليًا والمرتبطة بالموضوع في تكوين توقعاتهم وليس فقط الاعتماد على الخبرات الماضية[20].

باختصار، فإن فرض RE يعتمد على أن "التوقعات تكون رشيدة، بمعنى إدماج كل المعلومات المتاحة في الوقت الحالي لتكوين التوقع"[21] وليس فقط الاعتماد على المعلومات في الماضي.

النقد الموجه من أنصار RE إلى فرض AE تمت دراسته، واعتباره بشكل جيد، على الرغم من أن هناك العديد من الانتقادات التي وجهت أصلاً إلى فرض RE نفسه[22]. عمومًا هذا ليس المكان أو المجال الذي يتم فيه طرح تفاصيل هذه المواضيع. من الممكن القول كما قال Stephen Mc Nees التالي، "على أفضل تقدير، فرض التوقعات التكيفية من الممكن الدفاع عنه كفرض عملي مفوض عن صيغة أكثر تعقيدًا وقد تغير طريقة صياغة التوقع"[23].

مثال 8.17

إذا اعتبرنا النموذج الموجود في (11.4.17) المولود بطريقة التوقعات المتكيفة (مثل PPCE دالة في PPDI المتوقعة)، فإن γ معامل التوقعات ممكن الحصول عليه من (5.5.17) كالتالي:- $\gamma = 1 - 0.4106 = 0.5894$. وبالتالي باستكمال الاستعراض الذي تم لنموذج AE فمن الممكن أن نقول إن 59% من الاختلاف أو التعارض بين القيم الفعلية والمتوقعة قد تلاشى في خلال عام واحد.

---

(19) مثال نموذج Koyck، يمكن إثبات أنه في AE، توقعات المتغير هي متوسط أسي مرجح لقيم هذا المتغير في الماضي.

(20) لقراءات أكثر في فروض RE

G.K.show, op.cit., p.47

Steven M.Sheffrin, Ratioanl expectations, Cambrige. Press, New York, 1983 :- انظر

(21) Stephen K. McNees, "The Phillips Curve: Forward - or Backward-Looking?" New England Economic Review, July-August 1979, p. 50.

(22) لمزيد من التفاصيل عن فرض RE، انظر Micael C. Lovell, "Test of the Rational Expectations Hypotheris," Amerivcan Economic Reiew, March 1966, pp. 110-124.

(23) Stephen K. Mc Nees, op.cit., p.50.

6.17 أسلوب آخر رشيد لنموذج Koyck.. نموذج تعديل المخزون أو التعديل الجزيئي للمخزون :
ANOTHER RATIONALIZATION OF THE KOYCK MODEL: THE STOCK ADJUSTMENT, OR PARTIAL ADJUSTMENT, MODEL

نموذج التوقعات المتكيفة هو إحدى طرق ترشيد نموذج Koyck. أسلوب آخر رشيد قام به Marc Nerlove والمسمى بتعديل المخزون أو نموذج التعديل الجزيئي (PAM)[24]. لشرح هذا النموذج دعنا نعتبر النموذج المسرع المرن في النظرية الاقتصادية، والذي يفترض أن هناك توازنًا، وأمثلية أو قيمة طويلة المدى مرغوب فيها لمخزون المؤسسة والذي يعتبر ضروريًا لإنتاج كمية معينة من الناتج في إطار التكنولوچيا المتاحة ومعدل الفائدة وإلى ما غير ذلك من عوامل مؤثرة أخرى. للتبسيط دعنا نفترض أن الكمية المرغوبة من رأس المال هي Y_t^* والتي لها علاقة خطية مع الناتج X كالتالي:

$$Y_t^* = \beta_0 + \beta_1 X_t + u_t \qquad (1.6.17)$$

بما أن الكمية المرغوبة من رأس المال لايتم ملاحظتها مباشرة، قام Nerlove بافتراض التالي، وهو فرض معروف باسم التعديل الجزيئي أو تعديل المخزون كالتالي:

$$Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1}) \qquad (2.6.17)^{(25)}$$

بحيث إن δ، $0 < \delta \le 1$ معروفة باسم معامل التعديل و $Y_t - Y_{t-1}$ التغير الحقيقي و $Y_t^* - Y_{t-1}$ التغير المرغوب فيه. بما أن $Y_t - Y_{t-1}$ التغير في مخزون رأس المال بين فترتين ما هو إلا الاستثمار فإن (2.6.17) من الممكن كتابتها في شكل بديل كالتالي:

$$I_t = \delta(Y_t^* - Y_{t-1}) \qquad (3.6.17)$$

حيث إن I_t = الاستثمار في الفترة الزمنية t

(24) Marc Nerlove, Distributed lags and Demand analysis for Agricultural and other commodities, op.cit.

(25) بعض الكتاب لايضعون مقدار الخطأ العشوائي u_t إلى العلاقة (1.6.17) ولكن إضافة ذلك إلى العلاقة (2.6.17) معتقدين أنه إذا كان هناك بالفعل وضع توازي فلابد من إضافة مقدار الخطأ في حين أن طريقة التعديل ستكون ناقصة ولابد من إضافة مقدار الخطأ. لاحظ أن (2.6.17) من الممكن كتابتها كالتالي: $Y_t - Y_{t-1} = \delta(Y_{t-1}^* - Y_{t-1})$

معادلة (2.6.17) نفترض أن التغير الحقيقي في مخزون رأس المال (الاستثمار) في أي وقت زمني t هو نسبة y من التغير المطلوب في هذه الفترة. إذا كان δ=1 فإن ذلك يعني أن المخزون الفعلي يساوي المخزون المطلوب، بمعنى أن المخزون الفعلي يتعدل إلى المخزون المرغوب فيه بعد مرور بعض الوقت. عمومًا إذا كانت δ=0، فإن ذلك يعني أنه لم يحدث أي تغيير بما أن المخزون الفعلي عند الزمن t مساوي للقيمة المشاهدة في الفترة الزمنية السابقة. وبالفعل فإن δ من المتوقع لها أن تنحصر بين هذه القيم الطرفية، حيث إن التعديل للوصول إلى المخزون المرغوب فيه عادة سيكون غير كامل بسبب صرامة وقصور الشروط التعاقدية، ولذلك سمي النموذج بنموذج التعديل الجزيئي. لاحظ أن الطريقة التعديلية (2.6.17) ممكن كتابتها بشكل بديل كالتالي:

$$Y_t = \delta Y_t^* + (1 - \delta)Y_{t-1} \tag{4.6.17}$$

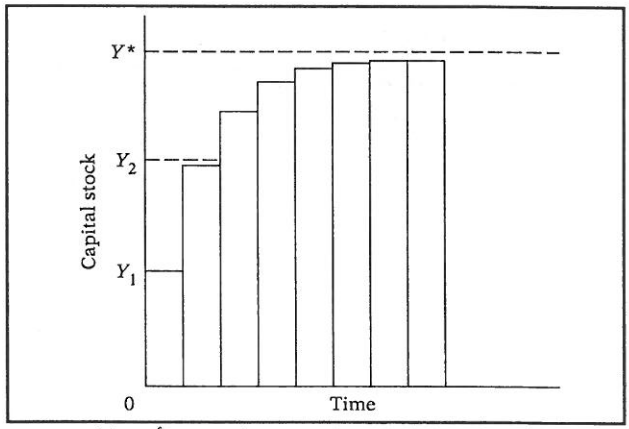
موضحاً أن مخزون رأس المال المشاهد في الزمن t هو متوسط مرجح لكل من مخزون رأس المال المرغوب فيه، ومخزون رأس المال الموجود في الفترة الزمنية السابقة، δ و (δ-1) يعتبران الأوزان المستخدمة. والآن بالتعويض بـ (1.6.17) في (4.6.17) نحصل على:

$$\begin{aligned} Y_t &= \delta(\beta_0 + \beta_1 X_t + u_t) + (1 - \delta)Y_{t-1} \\ &= \delta\beta_0 + \delta\beta_1 X_t + (1 - \delta)Y_{t-1} + \delta u_t \end{aligned} \tag{5.6.17}$$

هذا النموذج يسمى نموذج التعديلات الجزيئية (PAM) بما أن (1.6.17) تمثل طلبًا على مخزون رأس المال على المدى الطويل فإن (5.6.17) يمكن أن تسمى دالة الطلب على مخزون رأس المال في المدى القصير، حيث إنه في المدى القصير ليس بالضرورة أن يتساوى رأس المال الموجود مع قيمته في المدى البعيد. بمجرد تقدير دالة المدى القصير (5.6.17) والحصول على مقدر معامل التعديل δ (من معامل Y_{t-1}) فإننا نستطيع بسهولة اشتقاق دالة المدى البعيد بقسمة كل من $\delta\beta_0$ و $\delta\beta_1$ على δ وحذف قيمة Y في الفترة الزمنية المتأخرة والتي ستعطي في النهاية (1.6.17). هندسيًا نموذج التعديل الجزيئي ممثل في الشكل (6.17)[26]. وهذا الشكل Y^* هو مخزون رأس المال المرغوب فيه و Y_1 مخزون رأس المال الحقيقي الحالي. ولسهولة التوضيح تم افتراض δ = 0.5. هذا يعني أن المؤسسة تخطط لإغلاق نصف الفجوة في القيمة الفعلية والمرغوب فيها لمخزون رأس المال عند كل فترة.

---

(26) هذا مأخوذ من شكل (7.4) لكل من Rudiger Dornband و Strangly Fischer، Marcroeconomics, 3d ed., Mc craw- Hill, New York, 1984, p.216



شكل (6.17) التعديل التدريجي لمخزون رأس المال

وبالتالي في الفترة الأولى، تتحرك إلى $\frac{1}{2}$ باستثمار يساوي (Y_2-Y_1) والذي يساوي نصف (Y^*-Y_1). في كل فترة زمنية تابعة يتم إغلاق نصف الفجوة بين مخزون رأس المال عند بداية الفترة، ومخزون رأس المال المرغوب فيه Y^* نموذج التعديل الجزيئي يشابه كلاً من نموذج Koyck ونموذج التوقعات المتكيفة في أنه نموذج انحدار ذاتي. ولكن له مقدار خطأ أبسط، حيث إن مقدار الخطأ الأصلي u_t مضروب في ثابت δ. وضع في الاعتبار أنه على الرغم من التشابه في الشكل فإن نموذج التوقعات المتكيفة، ونموذج التعديلات الجزيئية مختلفين في المضمون. فالأول يعتمد على عدم التأكد (بالنسبة للأسعار في المستقبل، معدل الفائدة، ..) في حين الأخير يعتمد على الصلابة الفنية أو المؤسسية، وتكلفة معدل التغير وإلى ما غير ذلك من عوامل أخرى. عمومًا كل من هذين النموذجين نظريًا يعتبران أفضل من نموذج Koyck.

بما أن نموذج التوقعات المتكيفة، ونموذج التعديلات الجزيئية قريبان شكليًا، فإن المعامل γ المساوي 0.5894 لنموذج التوقعات المتكيفة من الممكن تفسيره كالمعامل δ الخاص بنموذج تعديل المخزون، بافتراض أن النموذج الأخير موجود في الحالة الحالية (بمعنى أن PPCE المتوقعة أو المرغوب فيها مرتبطة خطيًا مع PDPI الحالي). النقطة الرئيسية التي يجب وضعها في الاعتبار، أنه بما أن نموذج Koyck، ونموذج التوقعات المتكيفة، ونموذج تعديلات المخزون، باختلاف الفرق في شكل مقدار الخطأ، فإنهم جميعًا يعطون نفس النتيجة النهائية، وبالتالي لابد أن يكون الفرد شديد الحذر، عندما ينصح القارئ بأي هذه النماذج الأفضل في الاستخدام، ولماذا هي أفضل. وبالتالي فالباحثون لابد أن يحددوا الخلفية النظرية وراء النموذج المستخدم.

*7.17 الدمج بين نموذج التوقعات المتكيفة ونموذج التعديلات الجزيئية : COMBINATION OF ADAPTIVE EXPECTATIONS AND PARTIAL ADJUSTMENT MODELS

اعتبر النموذج التالي :

$$Y_t^* = \beta_0 + \beta_1 X_t^* + u_t \qquad (1.7.17)$$

حيث إن Y_t^* = مخزون رأس المال المرغوب فيه، و X_t^* مستوى الناتج المتوقع، بما أن كلاً من Y_t^* و X_t^* لايلاحظ أو يشاهد مباشرة، فمن الممكن استخدام أسلوب التعديل الجزيئي لـ Y_t^* ونموذج التوقعات المتكيفة لـ X_t^* فنصل إلى معادلة التقدير التالية (انظر تمرين 2.17) :

$$\begin{aligned} Y_t &= \beta_0\delta\gamma + \beta_1\delta\gamma X_t + [(1-\gamma)+(1-\delta)]Y_{t-1} \\ &\quad -(1-\delta)(1-\gamma)Y_{t-2} + [\delta u_t - \delta(1-\gamma)u_{t-1}] \\ &= \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + \alpha_3 Y_{t-2} + v_t \end{aligned} \qquad (2.7.17)$$

بحيث إن $v_t = \delta[u_t - (1-\gamma)u_{t-1}]$ هذا النموذج هو أيضًا نموذج انحدار ذاتي. الفرق الوحيد بينه وبين نموذج التوقعات الرياضية الخالصة، هو أن Y_{t-2} تظهر مع Y_{t-1} كمتغير مفسر. مثل Koyck ونماذج AE، مقدار الخطأ في (2.7.17) يتبع عملية المتوسط المتحرك. إحدى المميزات الأخرى في هذا النموذج أنه على الرغم من أن النموذج خطي في الـ α's إلا أنه غير خطي في المعاملات الأصلية.

التطبيق المهم لـ (1.7.17) هو فرض الدخل الدائم لـ Friedman والذي ينص على أن الدوام أو الاستهلاك في المدى البعيد هو دالة في الدخل في المدى البعيد أو الدخل الدائم [27].

تقدير (2.7.17) تظهر فيه نفس مشاكل التقدير مثل نموذج Koyck، ونموذج AE، حيث إن كل هذه النماذج تعاني من مشكلة الانحدار الذاتي، ولها نفس شكل الخطأ. بالإضافة إلى ذلك فإن (2.7.17) يتعلق بمشاكل التقدير غير الخطي التي نناقشها سريعًا في تمرين 10.17 وغير المذكورة بعمق في هذا الكتاب.

(*) اختياري.

(27) Milton Friedman, A Teory of Consumption Function, Princeton University Press, Princeton, N. J., 1957.

8.17 تقدير نماذج الانحدار الذاتي :
ESTIMATION OF AUTOREGRESSIVE MODELS

مما ناقشناه سابقًا ، فإن لدينا النماذج الثلاثة التالية :

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1}) \qquad (7.4.17)$$

ونموذج التوقعات المتكيفة :

$$Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma)Y_{t-1} + [u_t - (1-\gamma)u_{t-1}] \qquad (5.5.17)$$

ونموذج التعديلات الجزئية :

$$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta u_t \qquad (5.6.17)$$

كل هذه النماذج لها الشكل التالي المشترك :

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + v_t \qquad (1.8.17)$$

وبالتالي ، فهي نماذج انحدار ذاتي بطبيعتها . وبالتالي لابد من إلقاء نظرة على مشكلة التقدير في مثل هذه النماذج ، حيث إن طريقة المربعات الصغرى التقليدية لايمكن تطبيقها مباشرة في مثل هذه النماذج والسبب في ذلك : وجود متغيرات مفسرة عشوائية واحتمال وجود ارتباط تسلسلي .

والآن ، كما لاحظنا من قبل لتطبيق نظرية المربعات الصغرى التقليدية ، لابد من أن يكون المتغير المفسر العشوائي Y_{t-1} مستقل عن مقدار الخطأ v_t لتحديد ذلك لابد من التعرف على صفات v_t . إذا افترضنا أن مقدار الخطأ الأصلي u_t مستوفي كل الفروض التقليدية مثل $E(u_t) = 0$ و $var(u_t) = \sigma^2$ (فرض ثبات التباين) و$Cov(u_t)(u_{t+s}) = 0$ لكل $s \neq 0$ (فرض عدم وجود ارتباط ذاتي) ، فإن v_t قد لاتحقق ذلك كله . اعتبر على سبيل المثال مقدار الخطأ الموجود في نموذج Koyck والمساوي v_t $=(u_t - \lambda u_{t-1})$. بمعلومية الفروض الخاصة بـ u_t من الممكن بسهولة إثبات أن v_t مرتبط تسلسليًا كالتالي :

$$E(v_1 v_{t-1}) = -\lambda\sigma^2 \qquad [28](2.8.17)$$

(28) $E(v_t v_{t-1}) = E(u_t - \lambda u_{t-1})(u_{t-1} - \lambda u_{t-2})$

$= -\lambda E(u_{t-1})^2$ since covariances between u's are zero by assumption

$= -\lambda\sigma^2$

وهذا المقدار غير صفري (باستثناء عندما تكون λ صفراً). وبما أن Y_{t-1} تظهر في نموذج Koyck كمتغير مفسر، فمن الممكن أن تكون مرتبطة مع v_t (من خلال u_{t-1}). في الواقع من الممكن إثبات التالي:

$$\text{cov}[Y_{t-1}, (u_t - \lambda u_{t-1})] = -\lambda\sigma^2 \qquad (3.8.17)$$

وهو بالضبط مساوي لـ (2.8.17) ويمكن للقارئ أن يثبت نفس الشئ بالنسبة لنموذج التوقعات المتكيفة.

ماهي أهمية التوصل إلى أنه في نموذج Koyck، ونموذج التوقعات المتكيفة يكون المتغير المفسر العشوائي Y_{t-1} مرتبطاً مع مقدار الخطأ v_t؟ كما سبق وذكرنا، إذا كان المتغير المفسر في نموذج الانحدار مرتبطاً مع مقدار الخطأ العشوائي، فإن مقدرات الـ OLS ليست متحيزة فقط، وإنما أيضاً غير متسقة، بمعنى أنه حتى عندما يزداد حجم العينة، فإن المقدرات لاتؤول إلى قيم المجتمع الحقيقية[29]. وبالتالي فإن تقديرات نماذج Koyck، والتوقعات المتكيفة باستخدام طريقة OLS العادية سيؤدي إلى نتائج مغلوطة.

نموذج التعديل الجزئي يختلف، فعموماً في هذا النموذج يكون $v_t = \delta u_t$. بحيث إن $0 < \delta \leq 1$ وبالتالي إذا استوفى u_t الفروض الخاصة بنموذج الانحدار الخطي التقليدي المذكورة سابقاً، وكذلك سيكون δu_t. بالتالي فإن تقديرات OLS لنموذج التعديلات الجزئية ستكون مقدرات متسقة، على الرغم من أنها ستكون متحيزة (في العينات الصغيرة أو المحدودة)[30]. والسبب في الاتساق يرجع إلى أن: على الرغم من أن Y_{t-1} تعتمد على u_{t-1} وكل مقادير الأخطاء السابقة، فإنها لاترتبط مع مقدار الخطأ الحالي u_t.

(29) الإثبات خارج إطار هذا الكتاب ويمكن أن تجده في Gribiches, op.cit., pp.36-38 عموماً انظر الفصل (18) لخطوط عامة للإثبات في فكرة أخرى. انظر أيضاً:
Asatoshi Maeshivo, "Teaching regression with a lagged dependent variable and autocorrelated distribances", The Education, winter 1996, vol.27, no. 1, pp. 72-84.

(30) للإثبات انظر
J.Johnston, Econometric method, 3d ed., McGraw- Hill, New York, 1984, pp. 360-362.

انظر أيضاً
H.E. Doran and J. W. B. Guise, Single Equation Methods in Ecomometrics: Applied Regression Analysis, University of New England Teaching Monograph Series 3, Armidale, NSW, Australia, 1984, pp. 236-244.

وبالتالي بما أن u_t مستقلة تسلسليًا، فإن Y_{t-1} ستكون أيضًا مستقلة، أو على الأقل غير مرتبطة مع u_t، وبالتالي فإنها تحقق شرطًا مهمًا من شروط الـ OLS وهو شرط عدم الارتباط بين المتغيرات المفسرة، ومقدار الخطأ العشوائي. على الرغم من أن تقدير OLS لنموذج المخزون، أو نموذج التعديلات الجزيئية يعتبر تقديراً متسقًا بسبب بساطة تكوين مقدار الخطأ في مثل هذا النموذج، إلا أنه ليس من المفروض استخدام مثل هذا النموذج أكثر من نموذج Koyck أو نموذج التوقعات المتكيفة(31). ونحن ننصح القارئ بذلك بشدة، حيث إن اختيار نموذج ما لابد أن تكون له خلفية نظرية قوية، وليس فقط سهولة التقديرات الإحصائية التي تنتج عن استخدام مثل هذا النموذج. فكل نموذج يجب أن يتم استخدامه تبعًا للمميزات الخاصة به مع إعطاء اهتمام كبير لمقدار الخطأ العشوائي الخاص بالنموذج. إذا كان لايمكن تطبيق OLS مباشرة في نماذج Koyck، ونموذج التوقعات المتكيفة، فلابد من ابتكار طرق جديدة لحل مشكلة التقدير. وهناك العديد من طرق التقدير البديلة على الرغم من الصعوبة الحسابية لبعض منها. في الفقرة التالية دعنا نستعرض بعض هذه الطرق.

9.17 طريقة المتغيرات المساهمة (IV) :
THE METHOD OF INSTRUMENTAL VARIABLES (IV)

السبب الذي يجعل OLS لايمكن تطبيقها في حالة نموذج Koyck، أو نموذج التوقعات المتكيفة، هو أن المتغير المفسر Y_{t-1} يكون مرتبطًا مع مقدار الخطأ v_t. إذا تم بطريقة ما إلغاء مثل هذا الارتباط، فمن الممكن تطبيق OLS للحصول على مقدرات متسقة. كما ذكرنا من قبل (لاحظ أن: سيكون هناك مقدار من التحيز في العينات الصغيرة). كيف يمكن القيام بذلك؟ Liviatan اقترح الحل التالي(32).

دعنا نفترض أننا وجدنا متغيراً آخر مرتبطًا بشكل كبير من Y_{t-1}، ولكنه غير مرتبط مع v_t، ويكون نائبًا عن Y_{t-1} مع العلم أن v_t هو مقدار الخطأ الموجود في نموذج Koyck، أو نموذج التوقعات المتكيفة. مثل هذا النائب يسمى متغيراً مساهمًا(IV)(33).

---

(31) كما في : J.Johnstonnotes (op.cit., p.350), "[The] pattern of-adjustmit [sugget by the partial adjustment model].. may sometimes be implausible".

(32) N.Liviatan, "Consistent estimation of Distributed lags", International economic review, vol.4, January 1963, pp. 44-52.

(33) مثل هذه المتغيرات المساهمة تستخدم بكثرة في نماذج المعادلات التتابعية simulatoneous. (انظر الفصل 20).

Liviatan اقتراح استخدام X_{t-1} كمتغير مساهم عن المتغير Y_{t-1} وبعد ذلك تم اقتراح أن معاملات انحدار (1.8.17) من الممكن الحصول عليها عن طريق حل المعادلات الطبيعية التالية:

$$\begin{aligned}\sum Y_t &= n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{t-1}\\ \sum Y_t X_t &= \hat{\alpha}_0 \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum Y_{t-1} X_t \qquad (1.9.17)\\ \sum Y_t X_{t-1} &= \hat{\alpha}_0 \sum X_{t-1} + \hat{\alpha}_1 \sum X_t X_{t-1} + \hat{\alpha}_2 \sum Y_{t-1} X_{t-1}\end{aligned}$$

لاحظ أننا إذا كنا بصدد تطبيق OLS مباشراً للنموذج (1.8.17)، ستكون المعادلات الطبيعية العادية لـ OLS ستكون كالتالي (انظر فقرة 4.7):

$$\begin{aligned}\sum Y_t &= n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{t-1}\\ \sum Y_t X_t &= \hat{\alpha}_0 \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum Y_{t-1} X_t \qquad (2.9.17)\\ \sum Y_t Y_{t-1} &= \hat{\alpha}_0 \sum Y_{t-1} + \hat{\alpha}_1 \sum X_t Y_{t-1} + \hat{\alpha}_2 \sum Y_{t-1}^2\end{aligned}$$

الفرق بين المجموعتين من المعادلات الطبيعية يمكن ملاحظته بسهولة. Liviatan أثبت أيضاً أن α's المقدرة من (1.9.17) متسقة. في حين التقديرات التي نحصل عليها من (2.9.17) ممكن ألا تكون متسقة، حيث إن Y_{t-1} و $v_t\,[=u_t - \lambda u_t{-}1 \text{ or } u_t - (1-\gamma)u_{t-1}]$ قد يكونان مرتبطين في حين X_t و X_{t-1} غير مرتبطين مع v_t. (لماذا؟)

على الرغم من سهولة التطبيق العملي لهذه الطريقة، والمرتبطة بإيجاد متغير نائب أو مساهم مناسب، فإن طريقة Liviatan تعاني من مشكلة الارتباط الخطي المتعدد، حيث إن X_t و X_{t-1} الموجودين في المعادلات الطبيعية (1.9.17) سيكونان مرتبطين بشكل كبير.

(كما لاحظنا في الفصل (12)، معظم قيم المتغيرات الاقتصادية ذات السلاسل الزمنية عادة ما يكون بينها درجة عالية من الارتباط).

الخلاصة إذن أنه على الرغم من أن طريقة Liviatan تؤدي إلى الحصول على مقدرات متسقة، فإن هذه المقدرات غالباً ما تكون غير كافية[34].

(34) لترى كيف يمكن تحسين كفاءة التقدير. انظر في

Lawarence R.Kilen, A Textbook of Econometrics, 2d ed., Prentice-Hall, Englewood Cliffs, N.J., 1974, p.99.

وانظر أيضاً في :-

William H.Greene, Econometric Analysis, Macmillan, 2d ed., New York, 1993, pp. 535-538.

قبل الانتقال إلى نقطة أخرى. السؤال المهم هو: كيف يمكن إيجاد نائب أو متغير مساهم "جيد" للمتغير Y_{t-1}، بحيث يكون مرتبطًا مع Y_{t-1}، وغير مرتبط مع v_t؟ هناك بعض الاقتراحات في هذا الشأن، والتي نتعرض لها بشكل ما في التمارين (انظر تمرين 5.17) ولابد من القول بأن إيجاد متغير نائب أو مساهم جيد، ليست عملية سهلة، ولهذا السبب فإن طريقة IV لاتستخدم بكثرة في الواقع العملي، وقد يلجأ الباحث أكثر إلى طريقة الإمكان الأعظم للتقدير، والتي يعتبر شرحها خارج إطار هذا الكتاب[35].

هل هناك اختبار ما من الممكن استخدامه لاختبار المتغير المساهم؟ Dernis Sargan اقترح اختيار ما يعرف باسم اختيار SARG للقيام بمثل هذا الدور. الاختيار مشروح في الملحق A17، فقرة 1.A17.

10.17 اكتشاف الارتباط الذاتي في نماذج الانحدار الذاتي: اختبار DURBIN *h*
DETECTING AUTOCORRELATION IN AUTOREGRESSIVE MODELS: DURBIN *h* TEST

كما رأينا من قبل، فإن الارتباط التسلسلي المفترض في الخطأ v_t يجعل مشكلة التقدير في نماذج الانحدار الذاتي معقدة إلى حد ما: في نموذج تعديلات المخزون فإن مقدار الخطأ v_t ليس لديه ارتباط تسلسلي (من الدرجة الأولى) إذا كان مقدار الخطأ u_t في النموذج الأصلي غير مرتبط تسلسليًا، في حين في نموذج Koyck، ونموذج التوقعات المتكيفة، فإن v_t مرتبط تسلسليًا حتى إذا كان u_t مستقلاً تسلسليًا. السؤال إذن هو: كيف يمكن التعرف على أن الارتباط التسلسلي في مقدار الخطأ يظهر في نموذج الانحدار الذاتي؟

كما سبق وذكرنا في الفصل (12)، إحصاء Durbin- Watson *d* لايمكن استخدامه لمعرفة الارتباط التسلسلي (من الدرجة الأولى) في نماذج الانحدار الذاتي، حيث إن القيمة المحسوبة *d* في مثل هذه النماذج عادة ما تتجه إلى القيمة 2، وهي القيمة المتوقعة لـ *d* في التتابع الحقيقي العشوائي.

(35) لمزيد من التفاصيل عن طريقة ML، انظر في: J.Johnston, op. cit., pp. 366-371, كما في كل من App. 15A, App 4A

بمعنى آخر، إذا قمنا بحساب إحصاء d في مثل هذه النماذج، فإن هناك تحيزاً موجوداً أصلاً ضد التعرف على الارتباط التسلسلي (من الدرجة الأولى). رغم ذلك كله فإن العديد من الباحثين يلجأون لاستخدام القيمة d أفضل من أي شيء آخر. مؤخراً قام Durbin نفسه بتقديم اختيار جديد للعينات كبيرة الحجم للكشف عن الارتباط التسلسلي من الدرجة الأولى في نماذج الانحدار الذاتي(36). هذا الاختيار يسمى إحصاء h. تم تناول هذا الاختبار Durbon h من قبل في تمرين 36.12. للتسهيل فإن إحصاء h (مع تعديل بسيط في الترميز) هو:

$$h = \hat{\rho}\sqrt{\frac{n}{1 - n[\text{var}(\hat{\alpha}_2)]}} \qquad \textbf{(1.10.17)}$$

بحيث إن n هو حجم العينة، $\text{Var}(\hat{\alpha}_2)$ هو تباين معامل قيمة Y_t المتأخرة ($=Y_{t-1}$) في (1.8.17) و$\hat{\rho}$ تقدير للارتباط التسلسلي من الدرجة الأولى ρ، والذي ناقشناه من قبل في الفصل (12).

كما لاحظنا في تمرين 36.12، للعينات كبيرة الحجم، فإن Durbin أوضح أنه تحت صحة الفرض العدمي $\rho=0$، الإحصاء h الموجود في (1.10.17) يتبع التوزيع الطبيعي القياسي، بمعنى أن:

$$h_{\text{asy}} \sim N(0,1) \qquad \textbf{(2.10.17)}$$

بحيث إن asy تعني يؤول تقاربياً إلى.

في الواقع العملي، كما ذكرنا في الفصل (12)، ممكن تقدير ρ كالتالي:

$$\hat{\rho} \approx 1 - \frac{d}{2} \qquad \textbf{(3.10.17)}$$

ومن المثير للاهتمام، أنه على الرغم منه أنه لايمكن استخدام Durbin d لاختبار الارتباط الذاتي في نماذج الانحدار الذاتي، فإنه من الممكن استخدامه في حساب الإحصاء h.

دعنا نستعرض استخدام الإحصاء h بمثالنا 7.17.

في هذا المثال $n = 30$، $\hat{\rho} = (1 - \frac{d}{2}) = 0.4972$ (لاحظ أن: $d = 1.0056$)

و $\text{Var}(\hat{\alpha}_2) = \text{Var}(\text{PPCE}_{t-1}) = (0.1546)^2 = 0.0239$

(36) J.Durbin, "Testing for serial correction in least-squares regression when some of the regressors are lagged dependent variables", Econometrica, vol.38, 1970, pp.410- 421.

بوضع كل هذه القيم في (1.10.17) نحصل على:

$$h = 0.4972\sqrt{\frac{30}{1 - 30(0.0239)}} = 5.1191 \qquad (4.10.17)$$

وبما أن القيمة h تتبع التوزيع الطبيعي القياسي تحت صحة الفرض العدمي، فإن احتمال الحصول على هذه القيمة الكبيرة للـ h صغيرة جداً. تذكر أن احتمال أن قيمة التوزيع القياسي تزداد عند القيمة 3± هو احتمال صغير جداً. في المثال الحالي، نستنتج أن هناك ارتباطًا ذاتيًا موجبًا. بالطبع يجب أن نضع في الاعتبار أن الإحصاء h يؤول تقاربيًا إلى التوزيع الطبيعي القياسي، عينة المثال الحالي المكونة من 30 مفردة ليست بالضرورة تعتبر عينة كبيرة.

لاحظ الصفات التالية للإحصاء h :

1 - ليس من المهم على الإطلاق عدد المتغيرات المفسرة X، ولاعدد الفترات الزمنية المتأخرة Y المستخدمة في نموذج الانحدار لحساب h، نحتاج فقط إلى تباين معامل الفترة الزمنية المتأخرة Y_{t-1}.

2 - الاختبار لايجوز استخدامه إذا كان $[n \text{ Var}(\hat{\alpha}_2)]$ تزيد على 1 (لماذا؟). عمومًا لايحدث ذلك عادة.

3 - بما أن الاختبار خاص بالعينات ذات الحجم الكبير، فإن تطبيقه في حالة العينات الصغيرة غير مسموح به، كما هو موضح في Inder[37] و Kiviet[38]. اختبار Breuch- Godfery(BG)، والمعروف أيضًا باسم اختبار مضروب Lagrange، المشروح في الفصل (12)، هو اختبار له قوة إحصائية أكثر ليس فقط في حالة العينات ذات الحجم الكبير، ولكن أيضًا في العينات المحدودة، أو ذات الحجم الصغير، وهو بالتالي مفضل عن اختبار h[39].

(37) B.Inder, "An Approximation to the Null Distribution of the Durbin–Watson Statistic in Models Containing Lagged Dependent Variables," Econometric Theory, vol. 2, no. 3, 1986, pp. 413–428.

(38) J.F.Kiviet, "On the Vigour of Some Misspecification Tests for Modelling Dynamic Relationships," Review of Economic Studies, vol. 53, no. 173, 1986, pp. 241–262.

(39) Gabor Korosi, Laszlo Matyas, and Istvan P. Szekely, Practical Econometrics, Ashgate Publishing Compay, Brookfield, Vermont, 1992, p. 92.

11.17 مثال رقمي: الطلب على المال في كندا، I-1979 إلى IV-1988 : A Numerical Example: The demand for money in Canada, 1979-I to 1988-IV

لشرح استخدام النماذج التي تم استعراضها سابقًا، دعنا نستخدم المثال التطبيقي الخاص بالطلب على المال (المستوى الحقيقي للسيولة). بالتحديد دعنا نستعرض النموذج التالي[40]:

$$M_t^* = \beta_0 R_t^{\beta_1} Y_t^{\beta_2} e^{u_t} \qquad (1.11.17)$$

بحيث إن : M_t^* = الطلب على المال المرغوب فيه، أو في المدى البعيد (المستوى الحقيقي للسيولة).

R_t= معدل الفائدة طويل المدى، %

Y_t= الدخل القومي التجميعي الحقيقي.

لأسباب تتعلق بالتقدير الإحصائي، (1.11.17) يمكن التعبير عنها في شكل لوغاريتم الدالة كالتالي:

$$\frac{M_t}{M_{t-1}} = \left(\frac{M_t^*}{M_{t-1}}\right)^{\delta} \qquad 0 < \delta \le 1 \qquad (2.11.17)$$

معادلة (3.11.17) تنص على أن نسبة ثابتة (لماذا؟) من الفرق بين مستوى السيولة الحقيقي الموجود والمرغوب فيه يتم إنهاؤها في خلال فترة واحدة (عام واحد). في شكل اللوغاريتم، فإن معادلة (3.11.17) تأخذ الشكل التالي:

$$\ln M_t - \ln M_{t-1} = \delta(\ln M_t^* - \ln M_{t-1}) \qquad (3.11.17)$$

بالتعويض عن $\ln M_t^*$ من المعادلة (2.11.17) في المعادلة (4.11.17) وإعادة ترتيب حدود المعادلة نحصل على:

$$\ln M_t = \delta \ln \beta_0 + \beta_1 \delta \ln R_t + \beta_2 \delta \ln Y_t + (1-\delta) \ln M_{t-1} + \delta u_t \qquad (4.11.17)^{(41)}$$

---

(40) لنموذج مشابه انظر : Gregory C.Chow, "on the long-run and short run demand for money, Journal of political economy, vol.74, no.2, 1966, pp.111-131.

لاحظ أن إحدى مميزات الدالة المضروبة أن أس المتغيرات يعطي تقديراً مباشراً للمرونة (انظر الفصل 6).

(41) فيما سبق، لاحظ أن هذا النموذج هو بالضرورة غير خطي في المعلمات. وبالتالي على الرغم من أن OLS قد تعطي مقدرات غير متحيزة، مثلاً، $\beta_1\delta$ معاً قد لاتعطي مقدرات غير متحيزة لـ β_1 و δ منفردان خصوصاً إذا كان حجم العينة صغيراً.

والتي تسمى أحيانًا دالة الطلب قصيرة المدى على الأموال. (لماذا؟) لشرح دالة الطلب على الأموال السائلة الحقيقية في المدى القصير والطويل، دعنا نستعرض البيانات المعطاة في جدول (3.17). هذه البيانات الربع سنوية خاصة بكندا في الفترة من 1979 إلى 1988. المتغيرات تعرف كالتالي: M (كما هي معروفة بـ M1 عرض المال، الدولار الكندي (C$)، مليون)، P (مخفض السعر الضمني 100=1981)، GDP على أسعار ثابتة (C$، الملايين) و R (معدل الفائدة في 90 يومًا، %)[42]. M1 تم تخفيضها بـ P للحصول على أشكال مستويات السيولة الحقيقية.

مسبقًا، فإن الطلب الحقيقي على الأموال من المتوقع أن يكون مرتبطًا طرديًا (موجبًا) مع GDP (تأثير الدخل موجب) ومرتبط عكسيًا (سالبًا) مع R (كلما زاد معدل الفائدة كلما زاد احتمال الاحتفاظ بالأموال، حيث أموال M1 تدفع فائدة قليلة إذا دفعت فائدة أصلاً).

نتائج الانحدار كانت كالتالي[43]:

$$
\begin{aligned}
\widehat{\ln M_t} &= 0.8561 - 0.0634 \ln R_t - 0.0237 \ln \text{GDP}_t + 0.9607 \ln M_{t-1}\\
\text{se} &= (0.5101) \quad (0.0131) \quad (0.0366) \quad (0.0414)\\
t &= (1.6782) \quad (-4.8134) \quad (-0.6466) \quad (23.1972)\\
R^2 &= 0.9482 \quad d = 2.4582 \quad F = 213.7234 \qquad (5.11.17)
\end{aligned}
$$

---

(42) هذه البيانات تم الحصول عليها من B.Bhaskar Rao, ed., cointegration for the applied economist, st. Montin's press, New York, 1994, pp.210- 213.
البيانات الأصلية من 1956–I إلى 1988–IV ولكن لتسهيل الشرح، فقد بدأنا التحليل من الربع الأول في 1979.

(43) لاحظ هذه الصفة في الأخطاء القياسية المقدرة. الخطأ القياسي مثلاً لمعامل $\ln R_t$ يشير إلى الخطأ القياسي لـ $\widehat{\beta_1\delta}$ وهو مقدر لـ $\beta_1\delta$. لا توجد طريقة بسيطة للحصول على الخطأ القياسي لكل من $\hat{\beta}_1$، $\hat{\delta}$ منفردان من الخطأ القياسي لـ $\widehat{\beta_1\delta}$ خصوصًا عندما يكون حجم العينة صغيرًا نسبيًا. للعينات الكبيرة في الحجم، عمومًا فإن الأخطاء القياسية لكل من $\hat{\beta}_1$، $\hat{\delta}$ منفردان يمكن الحصول عليهما تقريبيين ولكن الحسابات تكون معقدة نوعًا ما. انظر في Jan Kementa, Elements of Econometrics, Macmillan, New York, 1971, p.444.

جدول (3.17) مستوى المال ، معدل الفائدة ، مؤشر الأسعار و GDP في كندا

Observation	M1	R	P	GDP
1979–1	22,175.00	11.13333	0.77947	334,800
1979–2	22,841.00	11.16667	0.80861	336,708
1979–3	23,461.00	11.80000	0.82649	340,096
1979–4	23,427.00	14.18333	0.84863	341,844
1980–1	23,811.00	14.38333	0.86693	342,776
1980–2	23,612.33	12.98333	0.88950	342,264
1980–3	24,543.00	10.71667	0.91553	340,716
1980–4	25,638.66	14.53333	0.93743	347,780
1981–1	25,316.00	17.13333	0.96523	354,836
1981–2	25,501.33	18.56667	0.98774	359,352
1981–3	25,382.33	21.01666	1.01314	356,152
1981–4	24,753.00	16.61665	1.03410	353,636
1982–1	25,094.33	15.35000	1.05743	349,568
1982–2	25,253.66	16.04999	1.07748	345,284
1982–3	24,936.66	14.31667	1.09666	343,028
1982–4	25,553.00	10.88333	1.11641	340,292
1983–1	26,755.33	9.616670	1.12303	346,072
1983–2	27,412.00	9.316670	1.13395	353,860
1983–3	28,403.33	9.333330	1.14721	359,544
1983–4	28,402.33	9.550000	1.16059	362,304
1984–1	28,715.66	10.08333	1.17117	368,280
1984–2	28,996.33	11.45000	1.17406	376,768
1984–3	28,479.33	12.45000	1.17795	381,016
1984–4	28,669.00	10.76667	1.18438	385,396
1985–1	29,018.66	10.51667	1.18990	390,240
1985–2	29,398.66	9.666670	1.20625	391,580
1985–3	30,203.66	9.033330	1.21492	396,384
1985–4	31,059.33	9.016670	1.21805	405,308
1986–1	30,745.33	11.03333	1.22408	405,680
1986–2	30,477.66	8.733330	1.22856	408,116
1986–3	31,563.66	8.466670	1.23916	409,160
1986–4	32,800.66	8.400000	1.25368	409,616
1987–1	33,958.33	7.250000	1.27117	416,484
1987–2	35,795.66	8.300000	1.28429	422,916
1987–3	35,878.66	9.300000	1.29599	429,980
1987–4	36,336.00	8.700000	1.31001	436,264
1988–1	36,480.33	8.616670	1.32325	440,592
1988–2	37,108.66	9.133330	1.33219	446,680
1988–3	38,423.00	10.05000	1.35065	450,328
1988–4	38,480.66	10.83333	1.36648	453,516

لاحظ أن : M1 = C$ بالملايين

P = مخفض الأسعار الضمني (100= 1981)

R = معدل الفائدة في 90 يوماً ، %

GDP= C$ بالملايين (أسعار 1981)

المصدر : Rao, op.cit., pp.210- 213

دالة الطلب قصيرة المدى المقدرة، تظهر أن مرونة الفائدة في المدى القصير لها الإشارة الصحيحة، وهي إحصائيًا معنوية، حيث إن قيمة P-value المرتبطة بها تقريبًا صفر. وبشكل غير متوقع، نجد أن مرونة الدخل في المدى القصير لها إشارة سالبة، على الرغم من أنها إحصائيًا لاتختلف عن الصفر. معامل التعديل δ يساوي 0.0393 = (1−0.9607) والذي يعني أن حوالي 4% من الفرق بين مستويات السيولة الحقيقية والمرغوب فيها يتم تسويته في مدة ربع سنة، مما يعتبر تسوية أو تعديلاً بطيئًا.

وبالعودة إلى المدى الطويل لدالة الطلب (2.11.17) كل مانحتاج إلى عمله هو الحصول على خارج قسمة دالة الطلب قصيرة المدى على δ (لماذا؟) ونسقط المقدار $\ln M_{t-1}$. النتائج هي كالتالي:

$$\widehat{\ln M_t^*} = 21.7888 - 1.6132 \ln R_t - 0.6030 \ln \text{GDP} \qquad \text{(6.11.17)}^{(44)}$$

وكما لاحظنا، فإن مرونة الفائدة طويلة المدى للطلب على المال هي أكبر من (مفهوم القيمة المطلقة) نظيرتها في المدى القصير، وذلك أيضًا صحيح بالنسبة لمرونة الدخل على الرغم من أنه في مثالنا الحالي، فإن المعنوية الاقتصادية والإحصائية لذلك تعتبر محل تساؤل.

لاحظ أن القيمة المقدرة Durbin- Watson d تساوي 2.4582، وبالتالي فهي قريبة من 2. وهذا يؤكد ملاحظتنا السابقة، أنه في نماذج الانحدار الذاتي قيمة d المحسوبة عادة ماتكون قريبة من 2. وبالتالي لايجب الثقة في قيمة d المحسوبة للكشف عن الارتباط التسلسلي (معرفة وجوده من عدمه) في بيانات المثال الحالي. حجم العينة في هذا المثال 40 مفردة، والذي يعتبر كبيرًا نوعًا ما ومناسبًا لتطبيق الاختبار h. في المثال الحالي، القارئ يمكنه أن يثبت أن قيمة h المقدرة هي 1.5008− وهي قيمة غير معنوية عند المستوى 5%، مما قد يقترح عدم وجود ارتباط ذاتي من الدرجة الأولى في مقدار الخطأ.

(44) لاحظ أننا لم نستعرض الأخطاء القياسية للمعاملات المقدرة لأسباب تم مناقشتها في الملاحظة رقم 43.

12.17 أمثلة توضيحية : ILLUSTRATIVE EXAMPLES

في هذه الفقرة، سنقدم بعض الأمثلة الخاصة بالنماذج الموزعة متأخراً لتوضيح كيف استخدم الباحثون مثل هذه النماذج في الدراسات التطبيقية (التجريبية).

مثال 9.17

معدل الفائدة الحقيقي والفيدرالي : The Fed and the real rate of interest

لتوضيح تأثير نمو M1 (العملة+ إيداع الشيكات) على مقياس معدل الفائدة الحقيقي للسندات Aaa قام كل من G. J. Santoni[45] و Courtenay C.Stone بتقدير النموذج الموزع مؤخراً كالتالي (معتمدين على بيانات شهرية) :

$$r_t = \text{constant} + \sum_{i=0}^{11} a_i \dot{M}_{t-i} + u_i \qquad (1.12.17)$$

بحيث إن r_t = مؤشر Moody للسندات Aaa، مطروح منه متوسط المعدل السنوي للتغير في مؤشر أسعار التعديل الموسمي للمستهلك خلال الـ 36 شهراً السابقة، والذي استخدم كمقياس لمعدل الفائدة الحقيقي و $\dot{M}_t$ = النمو الشهري في M_1.

وفقًا لمبدأ "حيادية مذهب المال" والذي ينص على أن المتغيرات الاقتصادية الحقيقية- مثل الإنتاج، العمالة، النمو الاقتصادي والمعدل الحقيقي للفائدة- لا تتأثر دائمًا بالنمو في المال، وبالتالي لاتتأثر بالضرورة بالسياسة المالية وفقًا لهذا المبدأ، فإن الاحتياطي الفيدرالي ليس له تأثير دائم على معدل الفائدة الحقيقي أيًا كان[46].

إذا كان هذا المذهب سليمًا، لابد من أن يتوقع الفرد أن المعاملات ai الخاصة بالنموذج الموزع متأخراً، وأيضًا مجاميعهم تكون إحصائيًا لا تختلف عن الصفر. لمعرفة ما إذا كان ذلك صحيحًا في المثال الحالي، فالباحثون قدروا (1.12.17) لفترتين زمنيتين مختلفتين. فبراير 1951 إلى سبتمبر 1979 وأكتوبر 1979 إلى نوفمبر 1982 . وقد وضعوا في اعتبارهم في الفترة الأخيرة التغيرات في سياسة المال الفيدرالية والتي منذ أكتوبر 1979 وضعت المزيد من الاهتمام على نمو عرض المال أكثر من معدل الفائدة، والذي كان سائداً في سياسات المال سابقًا. نتائج الانحدار الخاص بهم معروضة في جدول (4.17). النتائج يتضح تأييدها لمبدأ "حيادية مذهب المال" حيث إن في الفترة من فبراير 1951 إلى سبتمبر 1979 نجد نمو المال في الفترات الحالية، وأيضاً الفترات الزمنية المتأخرة لايوجد له أي معنوية إحصائية في التأثير على مقياس معدل الفائدة الحقيقي.

---

(45) The Fed and the real rate of interest", Review, Federal reserve bank of St. Louis, December 1982, pp. 8-18.

(46) The Fed and the Real Rate of Interest", Review, Federal Reserve Bank of St. Louis, December 1982, pp. 15.

في الفترة الأخرى أيضًا، تظهر صحة مبدأ حيادية مذهب المال ، حيث إن Eai لاتختلف إحصائيًا عن الصفر بخلاف المعامل a1 الذي يعتبر معاملاً معنويًا، ولكن له إشارة غير صحيحة . (لماذا؟)

جدول (4.17) تأثير النمو الشهري M1 على مقياس معدل الفائدة الحقيقي للسندات AAA : فبراير 1951 إلى نوفمبر 1982

$$r = \text{constant} + \sum_{i=0}^{11} a_i \dot{M}_{1_{t-1}}$$

	February 1951 to September 1979		October 1979 to November 1982	
	Coefficient	$\lvert t \rvert^*$	Coefficient	$\lvert t \rvert$
Constant	1.4885†	2.068	1.0360	0.801
a_0	−0.00088	0.388	0.00840	1.014
a_1	0.00171	0.510	0.03960†	3.419
a_2	0.00170	0.423	0.03112	2.003
a_3	0.00233	0.542	0.02719	1.502
a_4	−0.00249	0.553	0.00901	0.423
a_5	−0.00160	0.348	0.01940	0.863
a_6	0.00292	0.631	0.02411	1.056
a_7	0.00253	0.556	0.01446	0.666
a_8	0.00000	0.001	−0.00036	0.019
a_9	0.00074	0.181	−0.00499	0.301
a_{10}	0.00016	0.045	−0.01126	0.888
a_{11}	0.00025	0.107	−0.00178	0.211
$\sum a_i$	0.00737	0.221	0.1549	0.926
$\bar{R}^2$	0.9826		0.8662	
D-W	2.07		2.04	
RH01	1.27†	24.536	1.40†	9.838
RH02	−0.28†	5.410	−0.48†	3.373
NOB	344.		38.	
SER (= RSS)	0.1548		0.3899	

$\lvert t \rvert^*$ = قيمة t المطلقة

† تختلف معنويًا عن الصفر عند المستوى 0.05

المصدر : G.J.Sautoni and Courtenary C.stone. "The fed and the real of interest", review, federal reserve bank of st.louis, December 1982, p.16.

مثال 10.17

الاستهلاك التجميعي في المدى القصير والمدى الطويل في سيريلانكا ، 1967– 1993

The short- and long run aggregate consumption for sri-lanka, 1967- 1993.

افترض أن الاستهلاك C مرتبط خطيًا مع الدخل الدائم : X^*

$$C_t = \beta_1 + \beta_2 X_t^* + u_t \qquad (2.12.17)$$

بما أن X_t^* لاتلاحظ مباشرة، فإننا نحتاج إلى توصيف آلية أو طريقة لتوليد الدخل الدائم. افترض أننا اخترنا فرض التوقعات المتكيفة الموجود في (2.5.17). باستخدام (2.5.17) وتبسيطه نحصل على المعادلة المقدرة التالية : (cf 5.5.17)-

$$C_t = \alpha_1 + \alpha_2 X_t + \alpha_3 C_{t-1} + v_t \qquad (3.12.17)$$

حيث إن :

$$\alpha_1 = \gamma\beta_1$$
$$\alpha_2 = \gamma\beta_2$$
$$\alpha_3 = (1-\gamma)$$
$$v_t = [u_t - (1-\gamma)u_{t-1}]$$

كما نعلم β_2 يعطي متوسط الاستجابة للاستهلاك لكل زيادة 1$ في الدخل الدائم في حين أن α_2 يعطي متوسط الاستجابة للاستهلاك لكل زيادة 1$ في الدخل الحالي.

باستخدام بيانات سنوية لسيريلانكا في الفترة 1967–1993 والمعطاة في جدول (5.17) نحصل على نتائج الانحدار التالية[47] :

$$\hat{C} = 1038.403 + 0.4043X_t + 0.5009C_{t-1}$$
$$se = (2501.455) \quad (0.0919) \quad (0.1213) \qquad (4.12.17)$$
$$t = (0.4151) \quad (4.3979) \quad (4.1293)$$
$$R^2 = 0.9912 \quad d = 1.4162 \quad F = 1298.466$$

بحيث إن C= مصاريف الاستهلاك الخاصة و X =GDP كل منهما عند أسعار ثابتة. وقدمنا أيضًا معدل الفائدة الحقيقي في النموذج، ولكن غير معنوي إحصائيًا.

النتائج توضح أن الميل الحدي للاستهلاك في المدى القصير يساوي 0.4043 مما يعني أن الزيادة بوحدة نقدية واحدة في الدخل الحقيقي الحالي أو الملاحظ (مقاس بـ GDP الحقيقي) سيزيد متوسط الاستهلاك بحوالي 0.4 وحدة نقدية. ولكن إذا كانت الزيادة في الدخل مدعومة فإنه بالتالي الـ MPC الخارج من الدخل الدائم سيكون $\beta_2 = \gamma\beta_2/\gamma = 0.4.43/0.4991 = 0.8100$ أو حوالي 0.81 وحدة نقدية. بمعنى آخر، عندما يعدل المستهلك استهلاكه وفقًا للتغير في الدخل بوحدة نقدية واحدة، فإنه بمرور بعض الوقت سيزيد من استهلاكه بحوالي 0.81 وحدة نقدية.

(47) البيانات تم الحصول عليها من مكتب البيانات في :
Chandan Mukherjee, howard white and Marc Wuysts, Economtric and data analysis for developing countries, Routledge, New York, 1998.

البيانات الأصلية من جداول العالم لـ world bank's would tables

جدول (5.17) مصاريف الاستهلاك الخاص و GDP في سيريلانكا
Private consumption expenditature and GDP, Srilanka

Observation	PCON	GDP	Observation	PCON	GDP
1967	61,284	78,221	1981	120,477	152,846
1968	68,814	83,326	1982	133,868	164,318
1969	76,766	90,490	1983	148,004	172,414
1970	73,576	92,692	1984	149,735	178,433
1971	73,256	94,814	1985	155,200	185,753
1972	67,502	92,590	1986	154,165	192,059
1973	78,832	101,419	1987	155,445	191,288
1974	80,240	105,267	1988	157,199	196,055
1975	84,477	112,149	1989	158,576	202,477
1976	86,038	116,078	1990	169,238	223,225
1977	96,275	122,040	1991	179,001	233,231
1978	101,292	128,578	1992	183,687	242,762
1979	105,448	136,851	1993	198,273	259,555
1980	114,570	144,734			

لاحظ أن : PCON= مصاريف الاستهلاك الخاص

GDP= نمو الناتج المحلي

المصدر : انظر الملاحظة 47

والآن دعنا نفترض أن دالة الاستهلاك كالتالي :

$$C_t^* = \beta_1 + \beta_2 X_t + u_t \qquad (5.12.17)$$

في هذه المعادلة، فإن الاستهلاك في المدى البعيد أو الدائم C_t هو دالة خطية في الدخل الحالي أو الملاحظ، وبما أن C_t^* لاتلاحظ مباشرة، دعنا نستحضر نموذج التعديلات الجزئية (2.6.17). باستخدام هذا النموذج، وبعد عمل بعض العمليات الجبرية نحصل على :

$$\begin{aligned} C_t &= \delta\beta_1 + \delta\beta_2 X_t + (1-\delta)C_{t-1} + \delta u_t \\ &= \alpha_1 + \alpha_2 X_t + \alpha_3 C_{t-1} + v_t \end{aligned} \qquad (6.12.17)$$

من حيث الشكل، فإن هذا النموذج لايمكن تفريقه عن نموذج التوقعات المتكيفة (3.12.17). وبالتالي فإن نتائج الانحدار المعطاة في (4.12.17) ممكن تطبيقها هنا في هذا المثال. عمومًا هناك فرق رئيسي في تفسير النموذجين فبخلاف مشاكل التقدير الخاصة بالانحدار الذاتي وإمكانية وجود ارتباط تسلسلي في النموذج (3.12.17) فإن نموذج (5.12.17) هو نموذج لدالة الاستهلاك على المدى البعيد في حين (6.12.17) هو نموذج لدالة الاستهلاك في المدى القصير. β_2 تقيس الـ MPC في المدى البعيد، في حين ($\delta\beta_2$=)α_2 تعطي MPC في المدى القصير، الشكل التقليدي يمكن الحصول عليه من خلال القسمة على δ (معامل التعديل).

بالعودة إلى (4.12.17)، يمكن تفسير 0.4043 على أنه MPC في المدى القصير، وبما أن $\delta = 0.4991$ فإن MPC في المدى البعيد هو 0.81 . لاحظ أن معامل التعديل بحوالي 0.5 والذي يعني أنه عند أي فترة زمنية معطاة، فإن المستهلكين يعدلون فقط استهلاكهم بنصف المستوى في المدى البعيد أو المرغوب فيه.

في المثال يشير إلى نقطة مهمة، حيث إن الشكل الخارجي لنموذج التوقعات المتكيفة، ونموذج التعديلات الجزيئية أو نموذج Koyck متشابه جداً وبالنظر فقط إلى تقديرات الانحدار مثل الموجودة في (4.12.17) لايمكن الحكم بأي من هذه النماذج أفضل. ولهذا السبب يعتبر من الضروري تحديد الخلفية النظرية وراء اختيار نموذج ما لاستخدامه في التحليل العملي ثم يتبع ذلك عملية تطبيقية بشكل سليم.

وإذا كنا بصدد سلوك استهلاكي في شكل العادة، فإن نموذج التعديلات الجزيئية يكون مناسباً للاستخدام في حين إذا كان السلوك الاستهلاكي ينظر للمستقبل بمعنى أنه على أساس دخل مستقبلي متوقع، فإن نموذج التوقعات المتكيفة يكون هو النموذج المناسب. وإذا كنا في الحالة الأخيرة فإنه لابد من اعتبار مشكلة التقدير وإيجاد حل لها حتى نحصل على مقدرات متسقة. في الحالة العادية فإن OLS ستعطي مقدرات متسقة بافتراض صحة الفروض الرئيسية الطريقة OLS.

13.17 طريقة ALMON للنماذج الموزعة متأخراً: الفترات الزمنية الموزعة متأخراً المتعددة الحدود:

THE ALMON APPROACH TO DISTRIBUTED- LAG MODELS: THE ALMON OR POLYNOMIAL DISTRIBUTED LAG (PDL)(48):

على الرغم من الاستخدام الكثير لنموذج Koyck للفترات الزمنية المتأخرة في الواقع، فإن هذا النموذج مبني على افتراض أن معاملات β تتناقص هندسيًا مع طول الفترة الزمنية (انظر شكل 5.17). هذا الفرض يعتبر فرضًا مقيدًا جدًا للعديد من المواقف المختلفة. دعنا نعتبر على سبيل المثال الشكل (7.17).

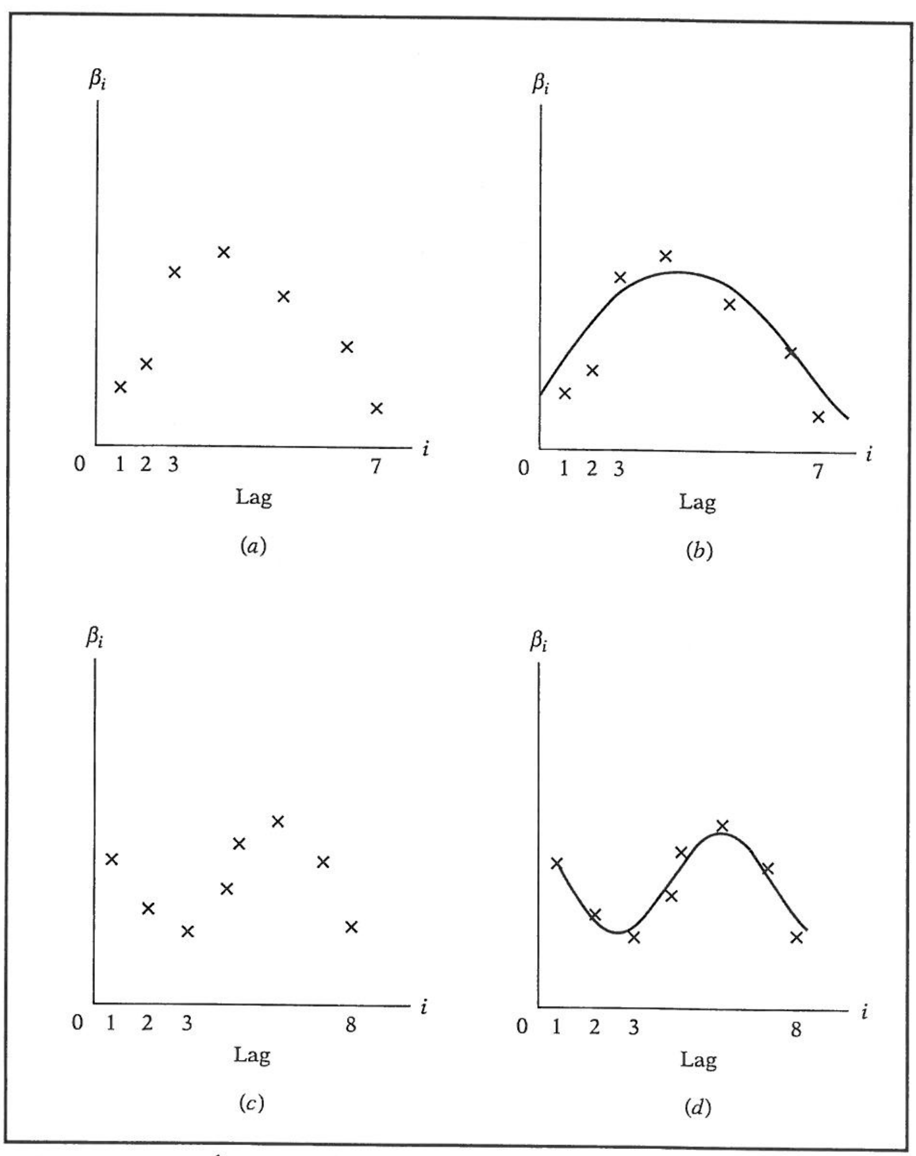
في شكل (a7.17) يفترض أن β's تتزايد في البداية ثم تتناقص، في حين شكل (c7.17) يفترض أن يتبع شكلاً دائريًا. وبالتالي فإن طريقة Koyck للنماذج الموزعة متأخرًا لن يناسب مثل هذه الحالات.

عمومًا بعد النظر إلى شكل (a7.17 و c) فإنه من الممكن التعبير عن β_i كدالة في i، طول الفترات الزمنية المتأخرة (الزمن)، ونطبق المنحنيات المناسبة للتعبير عن

(48) Sheirley Almon, "The Distributed Lag between Capital Appropriations and Expenditures", Econometrica, vol. 33. January 1965, pp. 178-196.

العلاقة الدالية بين المتغيرين، كما هو موضح في الشكل (b7.17 و d). هذه الطريقة تم اقتراحها عن طريق Shirley Almon. ولتوضيح هذا الأسلوب أكثر، دعنا نعود إلى النموذج الموزع متأخرًا المحدود والذي درسناه سابقًا وهو:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + u_t \qquad (2.1.17)$$



شكل (7.17) طريقة Almon للفترات الزمنية المتعددة المتأخرة

والذي يمكن كتابته أيضًا كالتالي :

$$Y_t = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + u_t \qquad (1.13.17)$$

وباتباع نظرية رياضية معروفة باسم نظرية Weierstrass ، فإن Almon افترض أن β_i يمكن تقريبها إلى متعددة محدودة من درجة مناسبة في i (طول الفترة الزمنية المتأخرة)(49) . فعلى سبيل المثال، إذا كانت طريقة الفترات الزمنية المتأخرة الموضحة في شكل (a7.17) متحققة فإنه يمكن كتابة التالي :

$$\beta_i = a_0 + a_1 i + a_2 i^2 \qquad (2.13.17)$$

والتي تعتبر معادلة تربيعية أو متعددة الحدود من الدرجة الثانية في i (انظر شكل b7.17) . عمومًا إذا كانت β's تتبع الشكل كالموجود في شكل (c7.17) يمكن كتابة التالي :

$$\beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3 \qquad (3.13.17)$$

والذي يعتبر متعددة حدود من الدرجة الثالثة في i (انظر شكل d7.17) وبشكل عام فإنه يمكن كتابة التالي :

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \cdots + a_m i^m \qquad (4.13.17)$$

والتي تعتبر متعددة حدود من الدرجة m في i . والتي تفترض أن m (درجة متعددة الحدود) تكون أقل من k (أطول فترة زمنية متأخرة) .

ولتوضيح كيفية استخدام طريقة Almon دعنا نفترض أن الـ β's تتبع الشكل الموجود في شكل (a7.17) وبالتالي متعددة حدود من الدرجة الثانية تكون هي الأنسب . بالتعويض عن (2.13.17) في (1.13.17) نحصل على :

$$\begin{aligned} Y_t &= \alpha + \sum_{i=0}^{k} (a_0 + a_1 i + a_2 i^2) X_{t-i} + u_t \\ &= \alpha + a_0 \sum_{i=0}^{k} X_{t-i} + a_1 \sum_{i=0}^{k} i X_{t-i} + a_2 \sum_{i=0}^{k} i^2 X_{t-i} + u_t \end{aligned} \qquad (5.13.17)$$

وبافتراض أن :

$$\begin{aligned} Z_{0t} &= \sum_{i=0}^{k} X_{t-i} \\ Z_{1t} &= \sum_{i=0}^{k} i X_{t-i} \\ Z_{2t} &= \sum_{i=0}^{k} i^2 X_{t-i} \end{aligned} \qquad (6.13.17)$$

(49) بشكل عام فإن النظرية تنص على أنه في فترة محدودة معلقة، فإن أي دالة متصلة يمكن تقريبها بشكل منتظم إلى متعددة حدود من درجة مناسبة .

وبالتالي فإنه يمكن كتابة (5.13.17) كالتالي:

$$Y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + u_t \qquad (7.13.17)$$

في طريقة Almon Y منحدرة على متغيرات منشأة Z، وليس على المتغيرات الأصلية X. لاحظ أن (7.13.17) ممكن تقديرها باستخدام طريقة OLS العادية تقديرات α و α_i سيكون لها نفس الخصائص الإحصائية الموجودة لمقدار الخطأ العشوائي u بما يضمن تحقق كل الشروط الخاصة بنموذج الانحدار الخطي التقليدي. في إطار ذلك، فإن أسلوب Almon له ميزة عن أسلوب Koyck حيث كما رأينا فإن الأخير له مشاكل خطيرة متعلقة بالتقدير، والتي تنتج من ظهور المتغير المفسر العشوائي Y_{t-1} والذي من المحتمل أن يكون مرتبطًا مع مقدار الخطأ (التشتت).

بمجرد تقدير الـ a's من (7.13.17) فإن الـ β 's الأصلية يمكن تقديرها من (2.13.17) [أو بشكل عام أكثر من (4.13.17)] كالتالي:

$$\begin{aligned}
\hat{\beta}_0 &= \hat{a}_0 \\
\hat{\beta}_1 &= \hat{a}_0 + \hat{a}_1 + \hat{a}_2 \\
\hat{\beta}_2 &= \hat{a}_0 + 2\hat{a}_1 + 4\hat{a}_2 \\
\hat{\beta}_3 &= \hat{a}_0 + 3\hat{a}_1 + 9\hat{a}_2 \\
&\cdots\cdots\cdots\cdots \\
\hat{\beta}_k &= \hat{a}_0 + k\hat{a}_1 + k^2\hat{a}_2
\end{aligned} \qquad (8.13.17)$$

قبل البدء في تطبيق أسلوب Almon، لابد من استعراض المشاكل التالية العملية:

1 - أكبر طول ممكن للفترات الزمنية المتأخرة k لابد أن يكون محدداً من قبل، وهنا من الممكن اتباع نصيحة Davidson, Mac Kinnon التالية:

أفضل أسلوب لعله يتضح من تحديد طول الفترات الزمنية المتأخرة أولاً، فالبدأ بقيمة كبيرة لـ q (طول الفترة الزمنية المتأخرة) ثم بعد ذلك نرى ما إذا كان توفيق النموذج جيدًا أم لا (معنوية التوفيق) مع تقليل هذه القيمة بدون وضع أي شروط مسبقة على شكل توزيع الفترات الزمنية المتأخرة(50).

هذه النصيحة مستوحاة من أسلوب الأعلى فالأقل لـ Hendry الذي ناقشناه في الفصل (13). تذكر أن هناك طولاً "حقيقيًا" للفترات الزمنية المتأخرة، واختيار

---

(50) Russell Davisdon and James G.Mac Kinnon, Estimation and inference in Econometric., Oxford university press, New York, 1993, 675-676.

عدد فترات زمنية قليل سيؤدى إلى "إهمال تحيز المتغيرات ذات الصلة بالموضوع" والذي ستكون عواقبه كما رأينا في الفصل (13)، وخيمة جداً. على الجانب الآخر، اختيار عدد فترات زمنية متأخرة أكثر من اللازم سيؤدي إلى "اشتمال تحيز متغيرات ليس لها صلة بالموضوع" وإن كانت عواقب ذلك أقل خطورة فالمقدرات يمكن تقديرها بطريقة OLS وستكون متسقة، ولكن التباين الخاص بالتقدير قد يكون أقل كفاءة.

من الممكن استخدام مقياس Akaike أو معلومة Schwarz التي تم استعراضها في الفصل (13) لاختيار الطول المناسب للفترات الزمنية المتأخرة. هذه المقاييس من الممكن أيضًا استخدامها لتحديد الدرجة المناسبة لمتعددة الحدود بالإضافة إلى المناقشة التالية في النقطة 2.

2 - بعد تحديد k، لابد أيضًا من تحديد درجة متعددة الحدود m بوجه عام، درجة متعددة الحدود يجب أن تكون على الأقل أزيد بواحد عن عدد نقاط التحول في المنحنى الذي يصف العلاقة بين β_i و i. وبالتالي في شكل (a7.17) هناك نقطة تحول واحدة، وبالتالي تستخدم متعددة حدود من الدرجة الثانية كتقريب جيد. في شكل (c7.17) هناك نقطتا تحول، وبالتالي متعددة حدود من الدرجة الثالثة ستعطي تقريبًا جيداً. مبدئيًا عمومًا من الممكن عدم معرفة عدد نقاط التحول، وبالتالي يكون اختيار m بشكل كبير يرجع للباحث. وعمومًا فإن النظرية تقترح شكلاً معينًا، في بعض الحالات في الواقع فإن الفرد يتمنى أن تكون متعددة الحدود من درجة قليلة (مثلاً 3 أو $m=3$) تعطي نتائج جيدة. بعد اختيار قيمة معينة للـ m إذا كنت تريد معرفة إذا ما كانت متعددة حدود من درجة أعلى، ستعطي توفيقًا أفضل، من الممكن اتباع التالي:

افترض أنك تريد الاختيار مابين متعددة حدود من الدرجة الثانية أو من الدرجة الثالثة. في متعددة الحدود من الدرجة الثانية تكون معادلة التقدير كالمعطاة في (7.13.17). أما في متعددة حدود من الدرجة الثالثة، فإن معادلة التقدير الخاص بها تكون كالتالي:

$$Y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + a_3 Z_{3t} + u_t \qquad \textbf{(9.13.17)}$$

بحيث إن $Z_{3t} = \sum_{i=0}^{k} i^3 X_{t-i}$. بعد الحصول على نتائج انحدار (9.13.17) إذا وجدت أن a_2 لها معنوية إحصائية، ولكن a_3 غير معنوية، فمن الممكن افتراض أن متعددة الحدود من الدرجة الثانية تعتبر منطقيًا تقريبًا جيداً.

على نحو آخر ، فإن Davidson و Mackinnon يقترحان "بعد تحديد q (طول الفترات الزمنية المتأخرة)، يمكن البدء في تحديد d (درجة متعددة الحدود) فمرة أخرى نبدأ بقيمة كبيرة ثم نقللها".

عموماً لابد من اعتبار مشكلة الارتباط الخطي المتعدد، والتي من المحتمل وجودها بسبب طريقة تكوين الـ Z's من X's كما هو موضح في (6.13.17) (انظر أيضاً (10.13.17)). كما سبق وذكرنا في الفصل (10) في حالات الارتباط الخطي المتعدد الشديدة $\hat{a}_3$ من الممكن أن تكون غير معنوية إحصائياً ليس بسبب أن قيمة a_3 الحقيقية مساوية للصفر، وإنما ببساطة لأن العينة المختارة لاتسمح لنا بتقييم أثر Z_3 على Y. وبالتالي في تعليقنا على النتائج، وقبل أن نقبل الاستنتاج الخاص بأن متعددة الحدود من الدرجة الثالثة ليست اختياراً جيداً، لابد أن نتأكد أولاً من أن مشكلة الارتباط الخطي المتعدد ليست شديدة الخطورة، ويمكن القيام بذلك بتطبيق الأساليب التي تمت مناقشتها في الفصل (10).

3 - بمجرد تحديد m و k يمكن تكوين الـ Z's، فعلى سبيل المثال ،إذا كانت $m = 2$ و $k = 5$ فإن Z's هي :

$$Z_{0t} = \sum_{i=0}^{5} X_{t-i} = (X_t + X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4} + X_{t-5})$$

$$Z_{1t} = \sum_{i=0}^{5} iX_{t-i} = (X_{t-1} + 2X_{t-2} + 3X_{t-3} + 4X_{t-4} + 5X_{t-5}) \quad \textbf{(10.13.17)}$$

$$Z_{2t} = \sum_{i=0}^{5} i^2 X_{t-i} = (X_{t-1} + 4X_{t-2} + 9X_{t-3} + 16X_{t-4} + 25X_{t-5})$$

لاحظ أن الـ Z's ما هي إلا توليفات خطية من الـ X's الأصلية، ولاحظ أيضاً السبب في أن Z's غالباً ما ستعاني من مشكلة الارتباط الخطي المتعدد.

قبل استعراض مثال رقمي، لاحظ مميزات استخدام طريقة Almon. أولاً إنها تعطي طريقة مرنة لاستخدام متغيرات في فترات زمنية متأخرة (انظر تمرين 17.17). أسلوب Koyck على الجانب الآخر محدود في ذلك، حيث إنه يفترض أن قيم الـ β's تتناقص هندسياً. ثانياً على عكس أسلوب Koyck فإن في طريقة ALmon لانقلق من وجود متغير تابع في فترة زمنية متأخرة كمتغير مفسر في النموذج، ولانقلق من المشاكل التابع لذلك في التقدير.

في النهاية إذا كانت هناك إمكانية من استخدام متعددة حدود من درجة قليلة، وتعطي توفيقاً جيداً فإن عدد المعاملات المطلوب تقديرها (الـ a's) سيكون أقل بشكل ملحوظ عن عدد المعاملات الأصلية (الـ β's).

دعنا الآن نستعرض المشاكل الخاصة بأسلوب Almon أولاً تحديد درجة متعددة الحدود، وأكبر قيمة ممكنة لعدد الفترات الزمنية المتأخرة يعتبر بوجه عام خاضعًا للباحث. ثانيًا لأسباب سابق ذكرها، فإن المتغيرات Z سيكون بها مشكلة ارتباط خطي متعدد. وبالتالي في نماذج مثل (9.13.17) فإن قيم a's المقدرة سيكون لها أخطاء قياسية كبيرة (منسوبة إلى قيم هذه التقديرات) وبالتالي سيكون واحد أو أكثر من هذه المعاملات غير معنوي إحصائيًا بسبب قيمة إحصاء t المرتبط به. ولكن هذا لايعني بالضرورة أن واحدًا أو أكثر من معاملات β's الأصلية لن يكون معنويًا إحصائيًا.

(إثبات هذه العبارة واضح بشكل ما ومقترح في تمرين 18.17)، وكنتيجة لذلك، فإن مشكلة الارتباط الخطي المتعدد يمكن أن تكون غير كبيرة كما يتصور البعض. إلى جانب ذلك كما نعرف، فإنه في حالات الارتباط الخطي المتعدد حتى إذا لم نكن قادرين على تقدير المعاملات بشكل دقيق، فإن توليفة خطية من هذه المعاملات (الدالة المقدرة) يمكن تقديرها بشكل دقيق.

مثال 11.17

شرح طريقة Almon في النموذج الموزع متأخرًا :

Illustration of the Almon distributed- lag model

لتوضيح طريقة Almon، جدول (6.17) يعطي بيانات عن قوائم الجرد Y والمبيعات X للولايات المتحدة الأمريكية في الفترة 1954–1999 لسهولة التوضيح، دعنا نفترض أن قيم قائمة الجرد تعتمد على المبيعات في السنة الحالية والسنوات الثلاث السابقة كالتالي :

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + u_t \qquad (11.13.17)$$

وبالتالي، نفترض أن الـ β's يمكن تقريبها باستخدام متعددة حدود من الدرجة الثانية كما هو موضح في (2.13.17). وبالتالي باتباع (5.13.17) يمكن كتابة.

$$Y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + u_t \qquad (12.13.17)$$

بحيث إن :

$$Z_{0t} = \sum_{i=0}^{3} X_{t-i} = (X_t + X_{t-1} + X_{t-2} + X_{t-3})$$

$$Z_{1t} = \sum_{i=0}^{3} i X_{t-i} = (X_{t-1} + 2X_{t-2} + 3X_{t-3}) \qquad (13.13.17)$$

$$Z_{2t} = \sum_{i=0}^{3} i^2 X_{t-i} = (X_{t-1} + 4X_{t-2} + 9X_{t-3})$$

متغيرات الـ Z يمكن تكوينها كما هو موضح في جدول (6.17). باستخدام البيانات الخاصة بـ y والـ Z's نحصل على الانحدار التالي:

$$\begin{aligned} \hat{Y}_t &= 25{,}845.06 + 1.1149Z_{0t} - 0.3713Z_{1t} - 0.0600Z_{2t} \\ se &= (6596.998) \quad (0.5381) \quad (1.3743) \quad (0.4549) \\ t &= (3.9177) \quad (2.0718) \quad (-0.2702) \quad (-0.1319) \\ R^2 &= 0.9755 \quad d = 0.1643 \quad F = 517.7656 \end{aligned} \tag{14.13.17}$$

لاحظ أن: بما أننا نستخدم عدد فترات زمنية متأخرة مساوي 3، فإن العدد الكلي للمشاهدات تم تقليله من 46 مفردة إلى 43 .

جدول (6.17) قوائم الجرد Y والمبيعات X ، الصناعة في الولايات المتحدة ، و Z's المكونه

Observation	Inventory	Sales	Z_0	Z_1	Z_2
1954	41,612	23,355	NA	NA	NA
1955	45,069	26,480	NA	NA	NA
1956	50,642	27,740	NA	NA	NA
1957	51,871	28,736	106,311	150,765	343,855
1958	50,203	27,248	110,204	163,656	378,016
1959	52,913	30,286	114,010	167,940	391,852
1960	53,786	30,878	117,148	170,990	397,902
1961	54,871	30,922	119,334	173,194	397,254
1962	58,172	33,358	125,444	183,536	427,008
1963	60,029	35,058	130,216	187,836	434,948
1964	63,410	37,331	136,669	194,540	446,788
1965	68,207	40,995	146,742	207,521	477,785
1966	77,986	44,870	158,254	220,831	505,841
1967	84,646	46,486	169,682	238,853	544,829
1968	90,560	50,229	182,580	259,211	594,921
1969	98,145	53,501	195,086	277,811	640,003
1970	101,599	52,805	203,021	293,417	672,791
1971	102,567	55,906	212,441	310,494	718,870
1972	108,121	63,027	225,239	322,019	748,635
1973	124,499	72,931	244,669	333,254	761,896
1974	157,625	84,790	276,654	366,703	828,193
1975	159,708	86,589	307,337	419,733	943,757
1976	174,636	98,797	343,107	474,962	1,082,128
1977	188,378	113,201	383,377	526,345	1,208,263
1978	211,691	126,905	425,492	570,562	1,287,690
1979	242,157	143,936	482,839	649,698	1,468,882
1980	265,215	154,391	538,433	737,349	1,670,365
1981	283,413	168,129	593,361	822,978	1,872,280
1982	311,852	163,351	629,807	908,719	2,081,117
1983	312,379	172,547	658,418	962,782	2,225,386
1984	339,516	190,682	694,709	1,003,636	2,339,112
1985	334,749	194,538	721,118	1,025,829	2,351,029
1986	322,654	194,657	752,424	1,093,543	2,510,189
1987	338,109	206,326	786,203	1,155,779	2,688,947

1988	369,374	224,619	820,140	1,179,254	2,735,796
1989	391,212	236,698	862,300	1,221,242	2,801,836
1990	405,073	242,686	910,329	1,304,914	2,992,108
1991	390,905	239,847	943,850	1,389,939	3,211,049
1992	382,510	250,394	969,625	1,435,313	3,340,873
1993	384,039	260,635	993,562	1,458,146	3,393,956
1994	404,877	279,002	1,029,878	1,480,964	3,420,834
1995	430,985	299,555	1,089,586	1,551,454	3,575,088
1996	436,729	309,622	1,148,814	1,639,464	3,761,278
1997	456,133	327,452	1,215,631	1,745,738	4,018,860
1998	466,798	337,687	1,274,316	1,845,361	4,261,935
1999	470,377	354,961	1,329,722	1,921,457	4,434,093

لاحظ أن : X، y وحدة قياسها المليون دولار ومعدلة موسميًا.

المصدر : economic report of the president, 2001, table B-57, p.340

الـ Z's مكونة كما هو مذكور في (13.13.17)

كتعليق مختصر على النتائج السابقة، نجد أنه من المتغيرات الثلاثة Z يوجد متغير واحد فقط Z_0 الذي له معنوية إحصائية منفرداً عند مستوى المعنوية 5%، ولكن الباقين غير معنويين، وبالتالي فإن قيمة F كبيرة جداً، مما يجعلنا نرفض الفرض العدمي، بمعنى أنه بشكل تجميعي، فإن الـ Z's ليس لها تأثير على Y . وقد يكون ذلك متوقعاً بسبب وجود مشكلة الارتباط الخطي المتعدد. ولاحظ أيضًا أن قيمة d المحسوبة صغيرة جداً. هذا لايعني بالضرورة أن البواقي تعاني من مشكلة الارتباط الذاتي، ولكن على الأرجح قيمة d المنخفضة تدل على أن النموذج المختار غير مناسب، وسنعلق على ذلك لاحقًا.

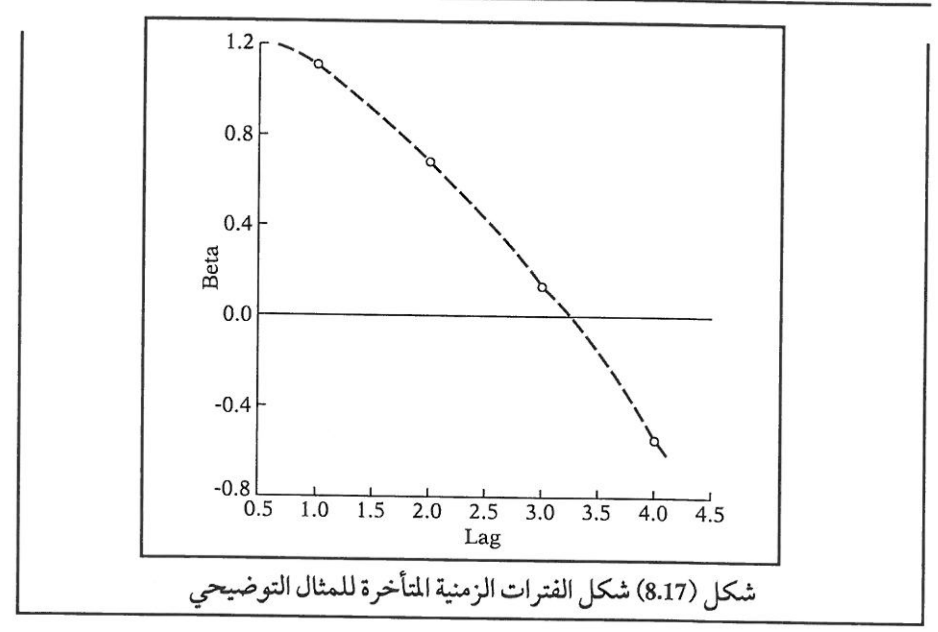
من خلال قيم a's المقدرة في (13.13.17) ممكن بسهولة تقدير βs الأصلية كما هو موضح في (8.13.17). في المثال الحالي، فإن النتائج كالتالي :

$$\begin{aligned}\hat{\beta}_0 &= \hat{a}_0 = 1.1149\\ \hat{\beta}_1 &= (\hat{a}_0 + \hat{a}_1 + \hat{a}_2) = 0.6836\\ \hat{\beta}_2 &= (\hat{a}_0 + 2\hat{a}_1 + 4\hat{a}_2) = 0.1321\\ \hat{\beta}_3 &= (\hat{a}_0 + 3\hat{a}_1 + 9\hat{a}_2) = -0.5394\end{aligned} \qquad (15.13.17)$$

وبالتالي فإن النموذج الموزع متأخراً المقدر الخاص بالمعادلة (11.13.17) هو كالتالي :

$$\begin{array}{lllllll}\hat{Y}_t = 25{,}845.0 & + \; 1.1150X_0 + & 0.6836X_{t-1} + & 0.1321X_{t-2} - & 0.5394X_{t-3} & \\ \text{se} = \; (6596.99) & (0.5381) & (0.4672) & (0.4656) & (0.5656) & (16.13.17)\\ t = \quad (3.9177) & (2.0718) & (1.4630) & (0.2837) & (-0.9537) & \end{array}$$

هندسيًا، قيم β_i المقدرة كالموضحة في شكل (8.17)



شكل (8.17) شكل الفترات الزمنية المتأخرة للمثال التوضيحي

المثال التوضيحي الحالي، يمكن الاستدلال به لاستعراض الصفات الإضافية الخاصة بطريقة الفترات الزمنية المتأخرة لـ Almon التالية:

1 - الأخطاء القياسية للمعاملات a يتم الحصول عليها مباشرة من انحدار OLS (14.13.17) ولكن الأخطاء القياسية لبعض معاملات الـ $\hat{\beta}$، التي يهتم الباحث بدراستها، لايمكن الحصول عليها مباشرة. ولكن يمكن الحصول عليها من الأخطاء القياسية الخاصة بالمعاملات a باستخدام معادلة إحصائية معروفة ومعطاة في التمرين (18.17). بالطبع لاتوجد حاجة لعمل تلك الحسابات باليد، حيث إن معظم حزم البرامج الإحصائية يمكنها القيام بذلك من خلال أكواد متوافرة بهذه الحزم. الأخطاء القياسية المعطاة في (15.13.17) تم الحصول عليها من Eviews 4.

2 - الـ $\hat{\beta}$'s التي تم الحصول عليها في (16.13.17) تسمى تقديرات غير مشروطة، بمعنى أنه لاتوجد قيود أو شروط مسبقة عليها. في بعض الحالات، عمومًا قد يحتاج الفرد إلى فرض مايسمى بقيود نقطة النهاية على الـ β's، حيث يتم افتراض أن β_0 و β_k (معامل الفترة الزمنية الحالية والفترة الزمنية المتأخرة k) يساويان الصفر. وبسبب بعض الأسباب النفسية والمؤسسية أو الفنية تكون قيمة المتغير المفسر في الفترة الزمنية الحالية ليس لها أي تأثير على القيمة الحالية للمتغير

المنحدر عليه، مما يعلل مساواة β_0 بالصفر. وبنفس المنطق، فإن قيمة معامل الفترة الزمنية المتأخرة k قد لايكون لها تأثير على المتغير المنحدر عليه، مما يؤيد فرض مساواة β_k بالصفر. في مثال قوائم الجرد معامل X_{t-3} له قيمة سالبة مما ليس له معنى اقتصادي. وبالتالي قد يلجأ الباحث إلى قيد مساواة هذا المعامل بالصفر[51]. بالطبع ليس بالضرورة وضع قيود على طرفي النهاية، فمن الممكن وضع القيد أو الشرط على المعامل الأول فقط، ويسمى ذلك القيد النهاية القريبة .أو يتم وضع القيد على المعامل الأخير فقط، ويسمى ذلك قيد النهاية القريبة .أو يتم وضع القيد على المعامل الأخير فقط، ويسمى ذلك قيد النهاية البعيدة. في مثال قوائم الجرد، تم شرح واستعراض ذلك في تمرين (28.17).

أحيانًا يتم تقدير β's وفقًا لشرط أن يساوي مجموعهم الصفر. ولكن لابد ألايتم استخدام هذا القيد بدون تفكير، حيث إن مثل هذا القيد يؤثر أيضًا على معاملات الفترة الزمنية المتأخرة الأخرى (غير المقيدة).

3 – بما أن اختيار عدد الفترات الزمنية المتأخرة، وأيضًا درجة متعددة الحدود المستخدمة يعتبر متروكًا نوعًا ما لحرية الباحث واختياره، فإن فكرة المحاولة والخطأ تكون حتمية، فلابد من استمرارية التدقيق في صحة الاختيار. وهنا قد يكون مفيداً استخدام أسلوب معلومات Schwarz و Akaike المشروح سابقًا في الفصل (13).

4 – بما أننا قدرنا (16.13.17) باستخدام ثلاث فترات زمنية متأخرة ومتعددة حدود من الدرجة الثانية، فإن ذلك نموذج مربعات صغرى مقيد. افترض أننا قررنا استخدام ثلاث فترات زمنية متأخرة، ولكن بدون استخدام طريقة متعددة الحدود لـ Almon. أي أننا قمنا بتقدير (11.13.17) باستخدام OLS. ماذا عن ذلك؟ دعنا أولاً نستعرض النتائج الخاصة بذلك:

$$\begin{aligned}
\hat{Y}_t &= 26{,}008.60 \quad + \; 0.9771X_t + \; 1.0139X_{t-1} - \; 0.2022\,X_{t-2} - \; 0.3935X_{t-3}\\
\text{se} &= \;(6691.12) \quad (0.6820) \quad (1.0920) \quad (1.1021) \quad (0.7186)\\
t &= \;(3.8870) \quad (1.4327) \quad (0.9284) \quad (-0.1835) \quad (-0.5476)\\
& R^2 = 0.9755 \qquad d = 0.1571 \qquad F = 379.51 \qquad \textbf{(17.13.17)}
\end{aligned}$$

(51) لتطبيق عملي، انظر D.B. Batten and Daniel thornton, “Polynomial Distributed Lags and the Estimation of the St. Louis Equation,” Review, Federal Reserve Bank of St. Louis, April 1983, pp. 13-52

إذا قارنت هذه النتائج مع النتائج الأخرى المعطاة في (16.13.17)، سنرى أن قيمة R^2 تقريبًا واحدة، على الرغم من أن شكل الفترات الزمنية المتأخرة في (17.13.17) يشير إلى وجود ارتفاعات وانخفاضات أكثر مما هو موجود في (16.13.17).

وكما هو ملاحظ من هذا المثال، لابد أن يكون الباحث حذرًا في استخدام طريقة Almon للفترات الزمنية الموزعة متأخرًا، حيث إن النتائج قد تكون حساسة لدرجة متعددة الحدود المختارة، أو لعدد معاملات الفترات الزمنية المتأخرة أو كلاهما معًا.

14.17 السببية في الاقتصاد.. اختبار GRANGER للسببية[52]:
CAUSALITY IN ECONOMICS THE GRANGER CAUSALITY TEST[52]

بالعود إلى فقرة 4.1، نلاحظ أنه على الرغم من أن تحليل الانحدار يتعامل مع اعتماد متغير واحد على عدد من المتغيرات، فإن ذلك لايعني بالضرورة السببية. بمعنى آخر وجود علاقة بين المتغيرات لاتثبت بالضرورة السببية أو اتجاه التأثير. ولكن في الانحدار المتعلق ببيانات السلاسل الزمنية، فإن الموقف يكون مختلفًا نوعًا ما، ونذكر هنا ما قاله أحد الكتاب وهو:

الزمن لايسير بالعكس. بمعنى أنه إذا وقع الحدث *A* قبل الحدث *B*، بالتالي فإنه من الممكن أن يكون *A* سببًا في *B*. عمومًا فإنه غير ممكن أن *B* سبباً لـ *A*. بمعنى آخر، الأحداث في الماضي يمكن أن تسبب الأحداث الحالية ولكن لايمكن القول بأن الأحداث في المستقبل هي سبب الأحداث الحالية[53]. (توضيح أكثر في أصل الكتاب) هذا بوجه عام يمثل فكرة الاختبار المسمى باختبار Granger للسببية[54].

ولكن يجب ملاحظة أن السؤال عن السببية هو سؤال فلسفي جدلي عميق، فعلى جانب شديد التطرف، يرى بعض الناس أن "كل شيء يسبب كل شيء" وعلى الجانب

---

(52) هناك اختبار آخر للسببية يستخدم أحيانًا يسمى اختبار Sims للسببية. سيتم التعرض له من خلال أحد التمارين.

(53) Gary Koop, Analysis of Economic Data, Johm Wiley& sons, New York, 2000, p.175.

(54) C.W.J.Granger, "Investigation causal relations by econometric models and cross spectral methods". Econometrica, July 1969, pp. 424-438.

على الرغم من أنه معروف باسم اختبار Granger للسببية، فمن الممكن أن يسمى أيضًا اختبار السببية لـ wiener-Granger حيث إنه يتم تناوله من قبل Wiener. انظر في Wiener, "The theory of prediction", in E.F. Beckenback, ed., Modern mathematics for engineers, Mc Graw-Hill, New York, 1956, pp.165-190.

(55) لمناقشة ممتازة في هذا الموضوع. انظر Arnold Zellner, "Causality and Econometrics," Camegie-Rochester Conference Series, 10, K. Brunner and A. H. Meltzer, eds., North Holland Publishing Company, Amsterdam, 1979, pp. 9–50.

شديد التطرف الآخر ينكر بعض الناس علاقة السببية أينما وجدت[55]. عالم الاقتصاد القياسي Edward Leamer يفضل استخدام اصطلاح الأسبقية عن السببية. Francis Diebold يفضل مصطلح السببية المتكهنة كما في كتابته التالية:

عبارة "y_i سبب y_j" تعتبر اختزالاً للعبارة الأدق، وإن كانت أطول وهي "yi تحتوي على معلومات تفيد في التنبؤ بـ y_i (من خلال مفهوم المربعات الصغرى)، بالإضافة إلى التاريخ السابق للمتغيرات الأخرى الموجودة في النموذج". وللاختصار نقول ببساطة y_i تسبب y_j[56].

اختبار Granger : The Granger Test

لشرح اختبار Granger، دعنا نستعرض السؤال الذي يطرح كثيراً في مجال الاقتصاد الكلي: هل GDP يسبب المعروض من المال M ($GDP \rightarrow M$) أو عرض المال M هو الذي يسبب GDP ($M \rightarrow GDP$) في حين أن السهم يشير إلى اتجاه العلاقة السببية. اختبار Granger للسببية يفترض أن المعلومات التي تساعد على التنبؤ بالمتغيرات محل الدراسة، GDP و M، موجودة فقط في بيانات السلاسل الزمنية الخاصة بهذه المتغيرات. الاختيار متعلق بتقدير الزوج التالي من الانحدارات:

$$\mathrm{GDP}_t = \sum_{i=1}^{n} \alpha_i M_{t-i} + \sum_{j=1}^{n} \beta_j \mathrm{GDP}_{t-j} + u_{1t} \qquad (1.14.17)$$

$$M_t = \sum_{i=1}^{n} \lambda_i M_{t-i} + \sum_{j=1}^{n} \delta_j \mathrm{GDP}_{t-j} + u_{2t} \qquad (2.14.17)$$

حيث يتم افتراض أن مقادير التشتت (الخطأ) u_{1t} و u_{2t} غير مرتبطين، فيما سبق، لاحظنا أنه بما أن لدينا متغيرين اثنين، فإننا نتعامل مع سببية ثنائية. في فصل الاقتصاد القياسي الخاص بالسلاسل الزمنية، سنقوم بتوسيع ذلك حتى يشتمل على السببية المتعددة المتغيرات من خلال أسلوب متجه الانحدار الذاتي (VAR).

معادلة (1.14.17) تفترض أن GDP الحالي مرتبط بقيم نفس المتغير في الماضي إلى جانب قيم المتغير M و (2.14.17) تفترض نفس السلوك بالنسبة لـ M. لاحظ أن هذه الانحدارات يمكن أن تجمع في أشكال النمو، GDP و M، في حين وجود نقطة فوق المتغير ستعني معدل نموه. والآن سنفرق بين أربع حالات:

(56) Francis X.Diebold, Elements of Forecasting, South Western Publishing, 2d ed., 2001. p. 254.

1 - السببية ذات الاتجاه الواحد من M إلى GDP، ستحدث إذا كانت القيم المقدرة لمعاملات الفترات الزمنية المتأخرة للـ M في (1.14.17) تختلف إحصائيًا عن الصفر كمجموعة واحدة (بمعنى أن $\Sigma\alpha_i \neq 0$) ومجموعة المعاملات المقدرة الخاصة بالمتغير GDP في الفترات الزمنية المتأخرة في (2.14.17) لا تختلف إحصائيًا عن الصفر (بمعنى أن $\Sigma\delta_j = 0$).

2 - على العكس، السببية ذات الاتجاه الواحد من GDP إلى M ستتواجد إذا كانت مجموعة معاملات M في الفترات الزمنية المتأخرة في (1.14.17) تختلف إحصائيًا عن الصفر (بمعنى أن $\Sigma\alpha_i = 0$) ومجموعة معاملات GDP في الفترات الزمنية المتأخرة في (2.14.17) تختلف إحصائيًا عن الصفر (بمعنى أن $\Sigma\delta_j \neq 0$).

3 - الاسترجاع أو ثنائية السببية تحدث عندما تكون مجموعة معاملات M و GDP تختلف إحصائيًا (معنوية) عن الصفر في كل من الانحدارين السابق ذكرهما.

4 - أخيرًا الاستقلال يحدث عندما تكون مجموعة معاملات M و GDP لاتختلف إحصائيًا (غير معنوية) عن الصفر في كل من الانحدارين السابق ذكرهما.

بوجه عام أكثر، بما أن المستقبل لايمكن أن يتنبأ بالماضي، إذا كان المتغير (Granger) X يسبب المتغير Y، فإن المتغيرات في X لابد أن تسبق التغيرات في Y. وبالتالي في الانحدار الخاص بـ Y على المتغيرات الأخرى (والتي تشتمل على قيم Y نفسها في فترات زمنية سابقة) إذا اشتمل الانحدار على قيم المتغير X السابقة أو في فترات زمنية متأخرة وأثرت إيجابيًا بشكل إحصائي على التقدير أو التنبؤ بـ Y، فإنه يمكن القول بأن X (Granger) تسبب Y. تعريف مماثل يمكن تطبيقه إذا كانت Y (Granger) تسبب X.

الخطوات التي يتم بها تحقيق اختبار Granger للسببية هي كالتالي. سنستعرض هذه الخطوات من خلال مثال المال- GDP المعطاة في المعادلة (1.14.17).

1 - قم بعمل انحدار لـ GDP الحالي على مقادير الـ GDP في الفترات الزمنية المتأخرة، وبعض المتغيرات الأخرى بخلاف متغيرات الـ M في الفترات الزمنية المتأخرة يجب عدم إدخالها في هذا الانحدار. وكما سبق وذكرنا في الفصل (8)، فإن ذلك الانحدار يعتبر انحدارًا مقيدًا. من هذا الانحدار أحصل على مجموع مربعات الأخطاء المقيد RSS_R.

2 - الآن قم بعمل الانحدار متضمنًا مقادير M في الفترات الزمنية المتأخرة، ووفقًا لمصطلحات الفصل (8)، هذا هو انحدار غير مقيد. من هذا الانحدار احصل على مجموع مربعات الأخطاء غير المقيدة RSS_{UR}.

3 - الفرض العدمي $H_0: \Sigma\alpha_i = 0$ يعني أن مقادير M في الفترات الزمنية المتأخرة لاتؤثر في هذا الانحدار.

4 - للقيام بعمل مثل هذا الاختبار، نطبق اختبار F الموجود في (9.7.8) وهو كالتالي:

$$F = \frac{(\mathrm{RSS}_R - \mathrm{RSS}_{\mathrm{UR}})/m}{\mathrm{RSS}_{\mathrm{UR}}/(n-k)} \qquad \textbf{(9.7.8)}$$

والذي يتبع توزيع F بدرجات حرية m و$(n-k)$. في المثال الحالي m تساوي عدد الفترات الزمنية المتأخرة للمقادير m و k هو عدد المعالم المقدرة في الانحدار غير المقيد.

5 - إذا كانت قيم F المحسوبة تزيد عن قيمة F الحرجة عند مستوى المعنوية المختار، نرفض الفرض العدمي، بمعنى أن قيم M في الفترات الزمنية المتأخرة تؤثر على هذا الانحدار. ويعتبر ذلك طريقة أخرى للقول بأن M تسبب GDP.

6 - الخطوات من 1 إلى 5 من الممكن تكرارها لاختبار النموذج (2.14.17) أي لاختبار ما إذا كانت GDP تسبب M.

قبل البدء في شرح اختبار Granger للسببية، هناك العديد من الأشياء التي يجب ملاحظتها أولاً وهي كالتالي:

1 - من المفترض أن المتغيرين الاثنين GDP و M ثابتان. ولقد ناقشنا مفهوم الثبات (عدم التغير) من خلال مصطلحات بديهية من قبل، وسنناقش ذلك بالتفصيل أكثر في الفصل (21) . أحيانًا بالتعامل مع الفروق الأولى للمتغيرات يجعلهم ثابتين إذا لم يكونا ثابتين في المستوى الأصلي.

2 - عدد الفترات الزمنية المتأخرة التي يتم إدخالها في اختبار السببية يعتبر من أهم الأسئلة العملية. كما في حالة النموذج الموزع متأخرًا فقد نحتاج إلى استخدام طريقة معلومات Schwarz و Akaike للقيام بمثل هذا الاختيار. ولكن يجب ملاحظة أن اتجاه السببية يعتمد بشكل كبير على عدد الفترات الزمنية المتأخرة المتضمنة في النموذج.

3 - قد افترضنا أن مقادير الأخطاء الموجودة في اختبار السببية غير مرتبطة. إذا لم يكن ذلك صحيحًا فإنه يمكن استخدام التحويلة المناسبة، كما سبق وشرحنا ذلك في الفصل 12[37].

4 - بما أن هدفنا هو اختبار السببية، فإنه يمكن للباحث ألا يستعرض القيم المقدرة للمعاملات الخاصة بالنموذج (1.14.17) و (2.14.17) بشكل تفصيلي (للاختصار) فقط نتائج اختبار F المعطاة في (9.7.8) ستكون كافية للغرض.

مثال 12.17

السببية بين المال والدخل : Casuality between Money and Income

R.W.Hafer استخدم اختبار Granger ليدرس طبيعة السببية بين GNP و M للولايات المتحدة في الفترة 1960-I to 1980-IV بدلاً من استخدام قيم النمو في هذه المتغيرات قام باستخدام معدلات النمو لكل من GNP و M وأيضًا استخدام أربع فترات زمنية متأخرة لكل من المتغيرات في الانحدارين اللذين سبق وتحدثنا عنهما من قبل. النتائج معطاة كالتالي[58]. الفرض العدمي في كل حالة أن المتغير محل الدراسة ليس هو السبب (كما في Granger) للمتغير الآخر.

Direction of causality	F value	Decision
$\dot{M} \rightarrow \dot{GNP}$	2.68	Reject
$\dot{GNP} \rightarrow \dot{M}$	0.56	Do not reject

هذه النتائج تعني أن اتجاه السببية يكون من نمو المال إلى نمو الـ GNP، حيث إن قيم F المقدرة معنوية عند المستوى 5%، قيمة F الحرجة هي 2.5 (لدرجات حرية 4 و 17). على الجانب الآخر لاتوجد علاقة سببية عكسية من نمو الـ GNP إلى نمو المال، حيث إن قيم F غير معنوية إحصائيًا.

مثال 17.3

السببية بين المال ومعدل الفائدة في كندا :
Causality between Money and Interest rate in Canada

بالعودة إلى بيانات كندا المعطاة في جدول (3.17). افترض أننا نريد معرفة ما إذا كانت هناك أي علاقات سببية بين عرض المال ومعدل الفائدة في كندا للفترة الربع سنوية

(57) لتفاصيل أكثر، انظر Wojciech W.Charemza and Derek F.Deadman, New Directions in Economeric Practice: General to Specific Modelling, Cointegration and Vector Autoregression, 3 d ed., Edward Elgar Publisher, 1997, Chap. 6.

(58) R.W.Hafer, "The Role of Fiscal Policy in the St. Louis Equation, "Review, Federal Resrve Bank of St. Louis, January 1982, pp. 17–22. See his footnote 12 for the details of the procedure.

بين 1979–1988 . لإظهار اعتماد اختبار Granger بشكل كبير على عدد الفترات الزمنية المتأخرة الموجودة في النموذج ، ستقدم نتائج اختبار F باستخدام عدد من الفترات الزمنية المتأخرة الربع سنوية في كل حالة ، الفرض العدمي هو أن معدل الفائدة لايسبب (Granger) المعروض من المال والعكس .

Direction of causality	Number of lags	F value	Decision
$R \rightarrow M$	2	12.92	Reject
$M \rightarrow R$	2	3.22	Reject
$R \rightarrow M$	4	5.59	Reject
$M \rightarrow R$	4	2.45	Reject (at 7%)
$R \rightarrow M$	6	3.5163	Reject
$M \rightarrow R$	6	2.71	Reject
$R \rightarrow M$	8	1.40	Do not reject
$M \rightarrow R$	8	1.62	Do not reject

لاحظ الخصائص التالية للنتائج السابقة لاختبار F: حتى الفترة الزمنية المتأخرة السادسة ، فإنه توجد علاقة سببية مزدوجة بين المعروض من المال ومعدل الفائدة ، بينما عند الفترة الزمنية المتأخرة الثامنة لاتوجد علاقة مماثلة إحصائية بين المتغيرين . مما يؤيد ماقيل سابقًا عن أن اختبار Granger حساس تجاه عدد المتغيرات الزمنية المتأخرة المستخدمة في النموذج .

مثال 14.17

السببية بين معدل نمو الـ GDP ومعدل الادخار الإجمالي في تسع بلاد في شرق آسيا :

Causality between GDP growth rate and Gross saving rate in nine East Asian countries

في دراسة عن السببية الثنائية بين معدل نمو GDP (g) ومعدل الادخار الإجمالي (S) النتائج معطاة في جدول (1.17)[59] .

بفرض المقارنة ، تم عرض نتائج مناظرة للولايات المتحدة الأمريكية في الجدول من معظم النتائج الموجودة في جدول (7.17) نجد أن غالبية دول شرق آسيا تكون السببية فيها من معدل نمو GDP في اتجاه معدل الادخار الإجمالي على العكس في الولايات المتحدة الأمريكية . وفي الفترة من 1950–1988 . وحتى الفترة الزمنية الثالثة المتأخرة السببية تحدث في الاتجاهين بين المتغيرين محل الدراسة ولكن في الفترة الزمنية المتأخرة الرابعة والخامسة السببية تحدث من معدل نمو الـ GDP في اتجاه معدل الادخار وليس العكس .

(59) هذه النتائج تم الحصول عليها من :

The East Asian Miracle: Economic crowth and public policy, published for the world bank by Oxford university press, 1993, p.224.

وكملخص لدراستنا عن اختبار السببية لـ Granger، يجب أن نضع في الاعتبار أن السؤال هو أننا نختبر ما إذا كان من الممكن التعرف إحصائيًا على اتجاه السببية عندما توجد علاقة بين المتغيرات مرتبطة بالفترات الزمنية المتأخرة بشكل مؤقت. إذا كانت هناك مثل هذه السببية فإنه من الممكن استخدام المتغير المسبب في تقدير للمتغير الآخر أفضل من استخدام قيم هذا المتغير الأخير في الماضي فقط.

في حالة اقتصاد دول شرق آسيا، يتضح أننا سنحصل على تقدير أفضل لمعدل الادخار الإجمالي باستخدام الفترات الزمنية المتأخرة لمعدل نمو الـ GDP عن استخدام الفترات الزمنية المتأخرة لمعدل الادخار الإجمالي فقط.

جدول (7.17) اختبارات السببية الثنائية لـ Granger بين معدل نمو الـ GDP ومعدل الادخار الإجمالي

Economy, years	Years of lags	Lagged right-hand side variable savings	Growth	Economy, years	Years of lags	Lagged right-hand side variable savings	Growth
Hong Kong, 1960–88	1	Sig	Sig	Philippines, 1950–88	1	NS	Sig
	2	Sig	Sig		2	NS	Sig
	3	Sig	Sig		3	NS	Sig
	4	Sig	Sig		4	NS	Sig
	5	Sig	Sig		5	NS	Sig
Indonesia, 1965	1	Sig	Sig	Singapore, 1960–88	1	NS	NS
	2	NS	Sig		2	NS	NS
	3	NS	Sig		3	NS	NS
	4	NS	Sig		4	Sig	NS
	5	NS	Sig		5	Sig	NS
Japan, 1950–88	1	NS	Sig	Taiwan, China, 1950–88	1	Sig	Sig
	2	NS	Sig		2	NS	Sig
	3	NS	Sig		3	NS	Sig
	4	NS	Sig		4	NS	Sig
	5	NS	Sig		5	NS	Sig
Korea, Rep. of, 1955–88	1	Sig	Sig	Thailand, 1950–88	1	NS	Sig
	2	NS	Sig		2	NS	Sig
	3	NS	Sig		3	NS	Sig
	4	NS	Sig		4	NS	Sig
	5	NS	Sig		5	NS	Sig
Malaysia, 1955–88	1	Sig	Sig	United States, 1950–88	1	Sig	Sig
	2	Sig	Sig		2	Sig	Sig
	3	NS	NS		3	Sig	Sig
	4	NS	NS		4	NS	Sig
	5	NS	Sig		5	NS	Sig

Sig : معنوية، NS : غير معنوية

لاحظ أن : النمو هو معدل نمو GDP عند الأسعار العالمية لـ 1985 .

المصدر : World bank, the East Asian Miracle

(*) ملاحظة على السببية وخارجية المنشأ : A Note on causality and Exogeneity *

كما سندرس في الفصول الخاصة بنموذج المعادلات في الجزء IV من هذا الكتاب، سنجد أن المتغيرات الاقتصادية عادة ما تقسم إلى طبقتين رئيسيتين، داخلية وخارجية. بوجه عام المتغيرات الداخلية هي المكافئة للمتغير التابع في نموذج انحدار

(*) اختياري

إحادي المعادلة ، والمتغيرات الخارجية هي المكافئة للمتغيرات X أو المنحدرة في هذا النموذج، مع افتراض أن متغيرات X غير مرتبطة مع مقدار الخطأ في هذه المعادلة(60) .

والآن دعنا نستعرض السؤال الشيق التالي: افترض أنه في اختبار السببية لـ Granger وجدنا أن المتغير X (Granger) يسبب المتغير Y بدون يسبب Y X (لاتوجد علاقة سببية مزدوجة). هل من الممكن في مثل هذه الحالة، أن نعامل X على أنه متغير خارجي؟ بمعنى آخر هل من الممكن استخدام سببي Granger (أو عدم السببية) لتحديد خارجية المنشأ للمتغير (أي تصنيف المتغير كمتغير خارجي)؟

للإجابة عن هذا السؤال ، لابد أن نفرق بين ثلاثة أنواع من خارجية المنشأ :
(1) ضعيف، (2) قوي، (3) ممتاز. لتبسيط الشرح، دعنا نفترض وجود متغيرين اثنين فقط Y و X، وافترض أيضًا أننا نقوم بانحدار Y_t على X_t، نقول إن X_t متغير خارجي ضعيف، إذا كان Y_t لايفسر X_t. في مثل هذه الحالة، تقدير واختبار نموذج الانحدار يمكن القيام به مشروطًا على قيم الـ X_t. وفي واقع الأمر، وبالعودة إلى الفصل (2). سندرك أن نموذج الانحدار مشروط على قيم المتغيرات المفسرة X. X_t يقال عنه متغير خارجي قوي إذا كانت قيم Y الحالية وقيم Y في فترات زمنية متأخرة لاتسره (بمعنى لاتوجد علاقة استرجاعية). ويقال عن X_t متغير خارجي ممتاز، إذا كانت محاولات Y و X لا تتغير حتى إذا تغيرت قيم X، بمعنى أن قيم المعاملات لاتتغير مع تغير قيم الـ X. إذا كان هذا هو الواقع فإن "نقد Lucas" الشهير قد يفقد قوته(61) .

السبب في التفرقة بين الأنواع الثلاثة الخاصة بخارجية المنشأ، هو أن "بوجه عام، خارجية المنشأ الضعيفة هي كل مايحتاج إليه للتقدير والاختبار، خارجية المنشأ القوية ضرورية للتنبؤ، وخارجية المنشأ الممتاز ضرورية لتحليل السياسات"(62) .

---

(60) بالطبع إذا كانت المتغيرات المفسرة تشتمل على واحد أو أكثر من المقادير الخاصة بالمتغير الداخلي في الفترات الزمنية المتأخرة، فإن هذا الشرط قد لايتحقق.

(61) Robert Lucas الحاصل على جائزة نوبل وضع من الآن فصاعداً افتراض أن العلاقة الموجودة بين المتغيرات الاقتصادية تتغير مع تغير السياسات. وفي مثل هذه الحالة المعاملات المقدرة من نموذج الانحدار ستكون قليلة القيمة من حيث التقدير والتنبؤ. لمزيد من التفاصيل انظر :
Oliver Blanchard, Macroeconomics, prentice hall, 1997, pp.371-372.

(62) Keith Cuthbertson, Stephen G.Hall and Mark P. Taylor, Applied Econometric Techniques, University of Michigan Press, 1992, P. 100.

بالعودة إلى سببية Granger، إذا كان هناك متغير، مثلاً Y لايسبب متغيراً آخر، مثلاً X، هل من الممكن إذن افتراض أن الأخير متغير خارجي؟ للأسف لاتوجد إجابة مباشرة على ذلك. إذا كنا بصدد خارجية المنشأ الضعيفة من الممكن إثبات أن سببية Granger ليست ضرورية ولا كافية لتوصيف المتغير بالمتغير الخارجي. على الجانب الآخر سببية Granger ضرورية (ولكن غير كافية) للخارجية القوية. إثبات هذه العبارات يقع خارج نطاق هذا الكتاب(63). في مجال دراسة هذا الكتاب، من الأفضل أن يتم الفصل بين مفهوم سببية Granger، ومفهوم خارجية المنشأ للمتغيرات، مع استخدام السابق كوسيلة وصفية مفيدة في تحليل بيانات السلاسل الزمنية. في الفصل (19) سنناقش اختبار ما لتحديد ما إذا كان المتغير متغيراً خارجيًا أم لا.

15.17 الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

1 - لأسباب نفسية، فنية أو مؤسسية، المتغير المنحدر عليه قد يستجيب للمتغيرات المنحدرة على فترات زمنية متأخرة. نماذج الانحدار التي تأخذ الفترات الزمنية المتأخرة للمتغير في الاعتبار معروفة باسم نماذج الانحدار ذات الفترات الزمنية المتأخرة، أو نماذج الانحدار المتحركة.

هناك نوعان من النماذج ذات الفترات الزمنية المتأخرة: الموزعة متأخراً، والانحدار الذاتي. في النموذج الأول، قيم المتغيرات المنحدرة الحالية، وقيم المتغيرات المنحدرة في الفترات الزمنية المتأخرة، هي متغيرات مفسرة. أما في النموذج الأخير، قيم المتغير المنحدر عليه في الفترات الزمنية المتأخرة تظهر كمتغير مفسر.

3 - النموذج الموزع متأخرًا الخالص من الممكن تقديره بـ OLS ولكن في مثل هذه الحالة هناك مشكلة الارتباط الخطي المتعدد، بما أن قيم المتغير المنحدرة في الفترات الزمنية المتأخرة تميل إلى الارتباط.

4 - وكنتيجة لذلك، بعض الطرق الأخرى تم استعراضها، وتشمل طريقة Koyck، التوقعات المتكيفة، وطريقة التعديلات الجزيئية. الطريقة الأولى طريقة ضعيفة في الخلفية الرياضية أما الاثنان الآخران فيقومان على أسس اقتصادية.

(63) لمناقشة بسيطة نسبيًا، انظر

G.S.Maddala, Introduction of econometrics, 2d ed., Macmillan, New York, 1992, pp. 394–395. and also David F. Hendry, Dynamic econometrics, Oxford University Press, New York, 1955, Chap. 5.

5 - ولكن كصفة مميزة لنماذج Koyck، التوقعات المتكيفة، ونماذج التعديلات الجزئية هي نماذج انحدار ذاتي بطبيعتها، حيث إن قيم المتغير المنحدرة عليه في فترات زمنية متأخرة تظهر كأحد المتغيرات المفسرة.

6 - الانحدار الذاتي، يطرح العديد من المشاكل في التقدير، إذا كان المتغير المنحدر عليه في الفترات الزمنية المتأخرة مرتبط مع مقدار الخطأ، فإن تقدير الـ OLS لمثل هذه النماذج ليست متحيزة فقط، وإنما أيضًا غير متسقة. التحيز وعدم الاتساق يحدث في حالة نموذج Koyck نموذج التوقعات المتكيفة، أما نموذج التعديلات الجزئية فهو مختلف، حيث إن تقديرات الـ OLS الخاص به تكون متسقة على الرغم من ظهور المتغير المنحدر عليه في فترات زمنية متأخرة.

7 - لتقدير نماذج Koyck، ونماذج التوقعات المتكيفة بشكل متسق، أشهر الطرق في ذلك هو طريقة المتغيرات المساهمة. المتغير المساهم وهو متغير مفوض عن المتغير المنحدر في فترة زمنية متأخرة، ولكن بخاصية عدم الارتباط مع مقدار الخطأ.

8 - كبديل عن نماذج الانحدار ذات الفترات الزمنية المتأخرة، تم استعراض نموذج Almon المتعدد الحدود الموزع متأخرًا، والذي لايعاني من مشاكل التقدير الخاصة بنماذج الانحدار الذاتي. المشكلة الرئيسية في طريقة Almon هي أنه يجب على الفرد بشكل شخصي، أن يحدد عدد الفترات الزمنية المتأخرة المستخدمة، ودرجة متعددة الحدود المستخدمة. وهناك طرق أساسية وغير أساسية لحل مشكلة الاختيار الخاصة بعدد الفترات الزمنية المتأخرة المستخدمة، ودرجة متعددة الحدود المستخدمة.

9 - وبغض النظر عن مشكلة التقدير والتي يمكن التغلب عليها، فإن نماذج الانحدار الذاتي الموزعة متأخرًا أثبتت فائدتها الشديدة في المجال الاقتصادي التطبيقي، حيث يحولون النظرية الاقتصادية الساكنة إلى أخرى متحركة من خلال اعتبارهم لعنصر الزمن. مثل هذه النماذج تساعدنا على التفرقة بين استجابة المتغير التابع لكل وحدة تغير في المتغير المفسر، أو المتغيرات المفسرة سواء على المدى القصير أو على المدى البعيد أيضًا. وبالتالي، فقد أثبتت مثل هذه النماذج فائدتها الشديدة لتقدير الأسعار، الدخل، التبديل والمرونة سواء في المدى الزمني القصير أو البعيد[64].

---

(64) للقراءة في تطبيقات مثل هذه النماذج. انظر
Arnold C.Harberger, ed., The demand for durable goods, university of Chicago press, Chocago, 1960.

10 - نظرًا لأن نماذج الانحدار الذاتي أو النماذج الموزعة متأخرًا تحتوي على الفترات الزمنية المتأخرة للمتغيرات، فقد برزت أهمية دراسة فكرة السببية في المتغيرات الاقتصادية وفي المجال التطبيقي، والنمذجة باستخدام سببية Granger حصلت على اهتمام العديد من الباحثين. ولكن يجب القول بأنه لابد من اختيار طريقة Granger بحذر شديد، حيث إنها طريقة شديدة الحساسية لطول الفترات الزمنية المتأخرة المستخدمة في النموذج.

11- حتى إذا كان المتغير (X) "المسبب- Granger" لمتغير آخر (Y)، فإن ذلك لايعني أن X متغير خارجي. فقد فرقنا بين ثلاثة أنواع لخارجية المنشأ: ضعيف، قوي، وممتاز- وقد أوضحنا أهمية هذه التفرقة.

تمــاريــن : EXERCISES

أسئلة : Questions

1.17 اشرح باختصار سبب صحة، أو خطأ أو عدم التأكد من العبارات التالية:

(a) كل نماذج الاقتصاد القياسي هي بالضرورة متحركة (دينامك).

(b) نموذج Koyck لن يكون ذا معنى إذا كان بعض معاملات المتغيرات الموزعة متأخرًا موجبة وبعضها سالبة.

(c) إذا تم استخدام OLS في تقدير نماذج Koyck ونماذج التوقعات المتكيفة، فإن التقديرات ستكون متحيزة ولكن متسقة.

(d) في نماذج التعديلات الجزيئية، مقدرات OLS تكون متحيزة إذا كان حجم العينة محدودًا.

(e) في حالة المتغيرات المنحدرة العشوائية، ومقادير الأخطاء ذات الارتباط الذاتي، فإن طريقة المتغيرات المساهمة ستعطي تقديرات غير متحيزة ومتسقة في نفس الوقت.

(f) في حالة وجود المتغير المنحدر عليه في فترة زمنية متأخرة كمتغير منحدر، إحصاء Durbin- Watson (d) لاختيار الارتباط الذاتي يعتبر عمليًا وسيلة فعالة لذلك.

(g) اختيار Durbin (h) يمكن استخدامه سواء في العينات صغيرة أو كبيرة الحجم.

(h) اختبار granger هو اختيار أسبقية أكثر من اختبار سببية.

2.17 أسس المعادلة (2.7.17).

3.17 اثبت المعادلة (3.8.17).

4.17 بافتراض أن الأسعار مكونة بناء على فرض التوقعات المتكيفة التالي:

$$P_t^* = \gamma P_{t-1} + (1-\gamma) P_{t-1}^*$$

حيث P^* هو السعر المتوقع، و P السعر الحقيقي.

أكمل الجدول التالي، بافتراض أن $\gamma = 0.5^*$.

Period	P*	P
$t-3$	100	110
$t-2$		125
$t-1$		155
t		185
$t+1$		—

5.17 اعتبر النموذج التالي:

$$M_t = \alpha + \beta_1 Y_t^* + \beta_2 R_t^* + u_t$$

حيث Y_{t-1} و v_t مرتبطان. لإزالة الارتباط، افترض أننا استخدمنا أسلوب المتغير المساهم. أولاً قم بانحدار yt على كل من X_{1t} و X_{2t} ثم احصل على قيم $\hat{Y}_t$ المقدرة من الانحدار، ثم قم بالانحدار التالي:

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 \hat{Y}_{t-1} + v_t$$

حيث إن: $\hat{Y}_{t-1}$ هي قيم Y المقدرة من الانحدار في المرحلة الأولى.

(a) كيف تزيل هذه الطريقة الارتباط بين Y_{t-1} و v_t في النموذج الأصلي؟

(b) ما هي مميزات هذه الطريقة الموصي بها عن طريقة Liviatan؟

6.17[†] **(a)** أسس (8.4.17).

(b) احسب وسيط الفترات الزمنية المتأخرة عندما تكون $\lambda = 0.2, 0.4, 0.6, 0.8$.

(c) هل توجد علاقة منتظمة مابين قيم الـ λ، وقيم وسيط الفترات الزمنية المتأخرة؟

(†) مختارة من G.K.Shaw, op.cit., p.26.

(*) اختياري

7.17 (a) اثبت أنه في نموذج Koyck، متوسط الفترات المتأخرة هو كالموجود في (10.4.17).

(b) إذا كانت λ كبيرة نسبيًا، ماذا يتضمن ذلك؟

8.17 باستخدام معادلة متوسط الفترات الزمنية المتأخرة المعطاة في (9.4.17) اثبت أن متوسط الفترات الزمنية المتأخرة يساوي 10.959 ربع سنوي المذكور في شرح جدول (1.17).

9.17 افترض $M_t = \alpha + \beta_1 Y_t^* + \beta_2 R_t^* + u_t$

بحيث إن: M = الطلب على مستوى السيولة الحقيقية، Y^*= الدخل الحقيقي المتوقع، R^*= معدل الفائدة المتوقع. افترض أن التوقعات مكونة كالتالي:

$$Y_t^* = \gamma_1 Y_t + (1 - \gamma_1) Y_{t-1}^*$$
$$R_t^* = \gamma_2 R_t + (1 - \gamma_2) R_{t-1}^*$$

بحيث إن γ_1 و γ_2 معاملات التوقع، كل منهما بين 0 و 1.

(a) كيف يمكنك التعبير عن M_t من خلال الكميات المشاهدة؟

(b) ما هي مشكلة التقدير التي تواجهك؟

10.17(*) إذا قدرت (2.7.17) بـ OLS. هل يمكنك اشتقاق مقدرات المعاملات الأصلية؟ ما هي المشكلة التي تواجهك؟ (للتفصيل، انظر +Roger N. Waud).

11.17 نموذج ارتباط تسلسلي.

اعتبر النموذج التالي :

$$Y_t = \alpha + \beta X_t + u_t$$

افترض أن u_t تتبع أسلوب الانحدار الذاتي لـ Markov من الدرجة الأولى والمعطى في الفصل (12)، أي أن:

$$u_t = u_{t-1} + \varepsilon_t$$

بحيث إن ρ هو معامل الارتباط الذاتي (من الدرجة الأولى)، و E_t مستوفاة كل الفروض الخاصة بطريقة OLS التقليدية. إذن كما موضح في الفصل (12)، النموذج:

(*) اختياري + Mis specification in the 'Partial Adjustment' and 'Adaptive Expectations' Models," International Economic Review, vol. 9, no. 2, June 1968, pp. 323–328.

$$Y_t = \alpha(1-\rho) + \beta(X_t - \rho X_{t-1}) + \rho Y_{t-1} + \varepsilon_t$$

سيحتوي على مقدار خطأ مستقل تسلسليًا، مما يجعل التقدير بـ OLS ممكنًا. ولكن هذا النموذج يسمى نموذج الارتباط التسلسلي، مشابه إلى حد كبير مع نماذج Koyck، التوقعات المتكيفة والتعديلات الجزيئية. كيف يمكنك أن تعرف في موقف ما، أي من النماذج السابق يعتبر مناسبًا(*)؟

12.17 اعتبر نموذج Koyck (أو لهذا الفرض التوقعات المتكيفة) المعطى في (7.4.17)، أي أن

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1})$$

افترض في النموذج الأصلي u_t تتبع أسلوب الانحدار الذاتي من الدرجة الأولى $u_t - \rho u_{1-t} = \varepsilon_t$ بحيث إن ρ هي معامل الارتباط الذاتي و ε_t مستوفاة كل الشروط التقليدية لـ OLS.

(a) إذا كانت $\rho = \lambda$. هل من الممكن تقدير نموذج Koyck بـ OLS؟

(b) هل التقديرات التي سيتم الحصول عليها متحيزة؟ متسقة؟ علل إجابتك.

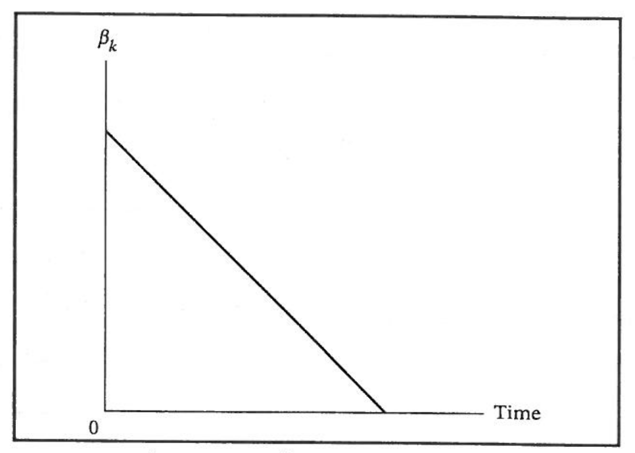
(c) إلى أي مدى يعتبر فرض $\rho = \lambda$ منطقيًا؟

13.17 النموذج الموزع متأخرًا الحسابي أو المثلثي(**).

هذا النموذج يفترض أن الباعث (المتغير المفسر) يبذل تأثيره الكبير في الفترة الزمنية الحالية، ثم ينخفض هذا التأثير بكميات متساوية حتى يصل إلى الصفر كلما زادت الفترة الزمنية المتأخرة بفترة واحدة في الماضي. هندسيًا ذلك موضح في شكل (9.17). باتباع التوزيع، افترض أننا نجري الانحدارات التالية:

(*) لشرح نموذج الارتباط التسلسلي انظر Zvi Griliches, "Distributed Lags: A survey," Econometrica, vol. 35, no. 1, January 1967, p. 34

(**) هذا النموذج مقترح من Irving Fisher في Note on a short-cut Method for Calculating Distributed Lags," International Statistical Bulletin, 1937, pp. 323–328.



شكل (9.17) طريقة الفترات الزمنية المتأخرة الحسابي أو المثلثي (Fisher's)

$$Y_t = \alpha + \beta\left(\frac{2X_t + X_{t-1}}{3}\right)$$

$$Y_t = \alpha + \beta\left(\frac{3X_t + 2X_{t-1} + X_{t-2}}{6}\right)$$

$$Y_t = \alpha + \beta\left(\frac{4X_t + 3X_{t-1} + 2X_{t-2} + X_{t-1}}{10}\right)$$

وهكذا، اختبار الانحدار الذي له أعلى R^2 " كأفضل " انحدار. علق على هذه الطريقة.

14.17 من البيانات الربع سنوية للفترة 1950–1960 .F.P.R.Brechling حصل على دالة الطلب التالية للعمالة في الاقتصاد الإنجليزي (الأرقام بين الأقواس تمثل الأخطاء القياسية)(*).

$$\widehat{\dot{E}}_t = 14.22 + 0.172Q_t - 0.028t - 0.0007t^2 - 0.297E_{t-1}$$
$$(2.61) \quad (0.014) \quad (0.015) \quad (0.0002) \quad (0.033)$$

$$\bar{R}^2 = 0.76 \qquad d = 1.37$$

حيث إن :
$\dot{E}_t = (E_t - E_{t-1})$

$Q =$ الناتج

$t =$ الزمن

(*) F.P.R.Brechling, "The Relationship between Output and Employment inBritish Manufacturing Industries," Review of Economic Studies, vol. 32, July 1965.

المعادلة السابقة تعتمد على افتراض أن المستوى المرغوب فيه من التوظيف E_t^* هو دالة في الناتج، الزمن ومربع الزمن وأيضًا على الفرض $E_t - E_{t-1} = \delta (E_t^* - E_{t-1})$ بحيث إن δ هو معامل التعديل ويقع بين 0 و 1.

(a) فسر الانحدار السابق.

(b) ما هي قيمة δ؟

(c) اشتق دالة الطلب على العمالة في المدى البعيد من دالة الطلب المقدرة في المدى القصير.

(d) كيف يمكنك اختبار الارتباط التسلسلي في النموذج السابق؟

15.17 في دراسة عن طلب السوق على الجرارات Griliches استخدم النموذج التالي(*):

$$T_t^* = \alpha X_{1,t-1}^{\beta_1} X_{2,t-1}^{\beta_2}$$

بحيث إن: T^* = المخزون المرغوب فيه من الجرارات.

X_1 = الأسعار النسبية للجرارات.

X_2 = معدل الفائدة.

باستخدام نموذج تعديلات المخزون، حصل على النتائج التالية عن الفترة 1921–1957:

$$\widehat{\log T_t} = \text{constant} - \underset{(0.051)}{0.218} \log X_{1,t-1} - \underset{(0.170)}{0.855} \log X_{2,t-1} + \underset{(0.035)}{0.864} \log T_{t-1}$$

$$R^2 = 0.987$$

بحيث إن الأرقام بين الأقواس هي الأخطاء القياسية المقدرة.

(a) ما هي القيمة المقدرة لمعامل التعديل؟

(b) ما هي مرونة السعر في المدى القصير والبعيد؟

(c) ما هي مرونة الفائدة المرتبطة بها؟

(d) ما هي أسباب ارتفاع أو انخفاض معدل التعديل في النموذج الحالي؟

(*) Zvi Griliches, "The demand for a Durable Input: Farm Tractors in the United States, 1921–1957," in Arnold C. Harberger, ed., The Demand for Durable Goods, University of Chicago Press, Chocago, 1960.

16.17 عندما يظهر المتغير التابع في فترة زمنية متأخرة كمتغير مفسر فإن R^2 عادة تكون أكبر من قيمتها في حالة عدم حدوث ذلك. ما هي أسباب ذلك؟

17.17 اعتبر شكل الفترات الزمنية المتأخرة المعطى في شكل (10.17). ما هي درجة متعددة الحدود التي تختارها لشكل الفترات الزمنية المتأخرة ولماذا؟

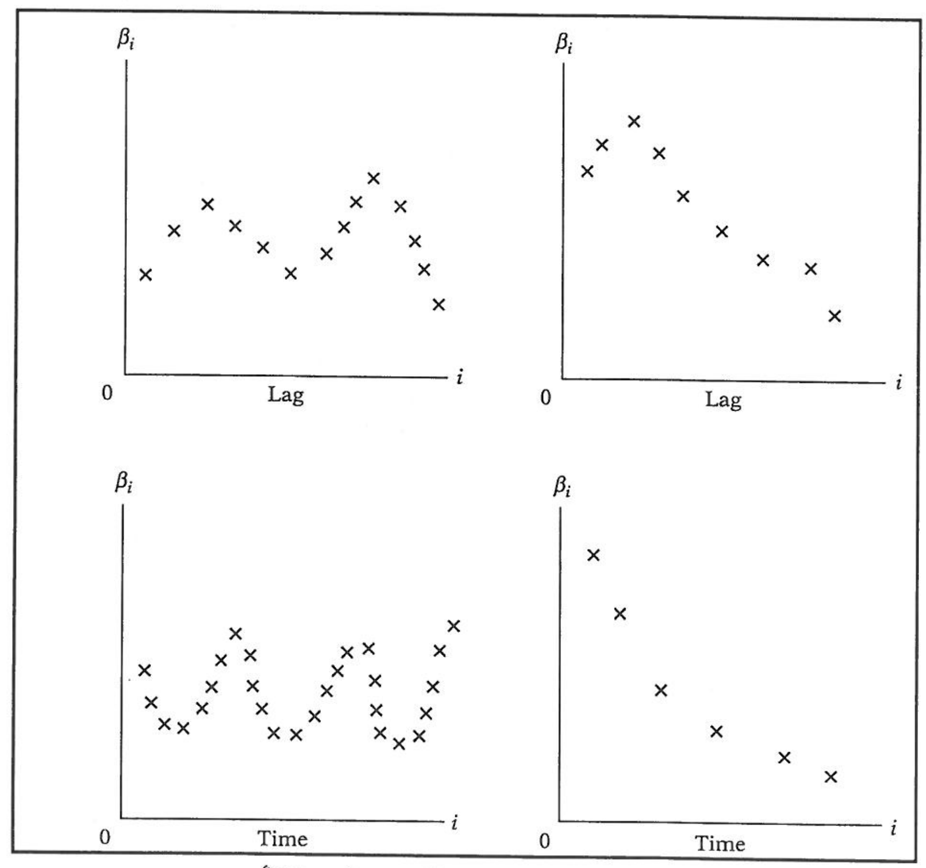
18.17 اعتبر المعادلة (4.13.17).

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \cdots + a_m i^m$$

للحصول على تباين $\hat{\beta}_1$ من تباين $\hat{\alpha}_i$ تستخدم الصيغة التالية:

$$\text{var}(\hat{\beta}_i) = \text{var}(\hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2 + \cdots + \hat{a}_m i^m)$$

$$= \sum_{j=0}^{m} i^{2j}\,\text{var}(\hat{a}_j) + 2\sum_{j<p} i^{(j+p)}\text{cov}(\hat{a}_j \hat{a}_p)$$



شكل (10.17) شكل مفترض للفترات الزمنية المتأخرة

(a) باستخدام الصيغة السابقة، أوجد تباين $\hat{\beta}_i$ معبر عنه كالتالي:

$$\hat{\beta}_i = \hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2$$

$$\hat{\beta}_i = \hat{a}_0 + \hat{a}_1 i + \hat{a}_2 i^2 + \hat{a}_3 i^3$$

(b) إذا كان تباين $\hat{a}_i$ كبيرًا مقارنًا بقيمة $\hat{a}_i$، هل سيكون تباين $\hat{\beta}_i$ كبيرًا أيضًا؟ علل إجابتك.

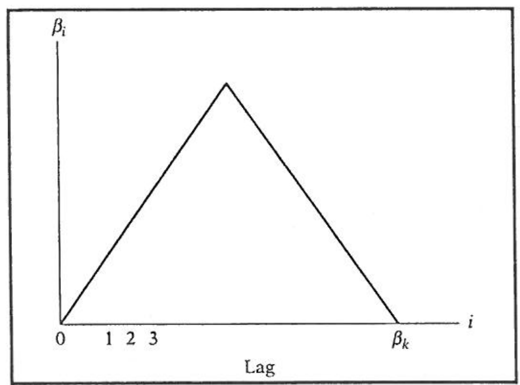
19.17 اعتبر النموذج الموزع متأخرًا التالي:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + u_t$$

افترض أن β_i من الممكن التعبير عنها على نحو كاف بمتعددة حدود من الدرجة الثانية كالتالي:

$$\beta_i = a_0 + a_1 i + a_2 i^2$$

كيف يمكنك تقدير الـ β's إذا كنت تريد فرض القيد التالي $\beta_0 = \beta_4 = 0$؟



شكل (11.17) مقلوب V للنموذج الموزع متأخرًا

20.17 مقلوب V للنموذج الموزع متأخرًا. اعتبر النموذج الموزع متأخرًا المحدود بالفترة k التالي:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + u_t$$

F.Deleeuw اقترح شكل الـ β's الموجودة في الشكل (11.17) حيث إن الـ β's تتبع مقلوب شكل V. للتبسيط دعنا نفترض أن k (أكبر طول للفترة الزمنية المتأخرة) هي عدد زوجي، وافترض أيضًا أن β_0 و β_K مساويان للصفر. Deleeuw اقترح الطريقة التالية للـ βs(*):

(*) انظر لمقال: "The Demand for Capital Goods by Manufacturers: A Study of Quarterly Tim Serirs," Econometrica, vol. 30, no. 3, July 1962, pp. 407–423.

$$\beta_i = i\beta \qquad 0 \le i \le \frac{k}{2}$$
$$= (k-i)\beta \qquad \frac{k}{2} \le i < k$$

كيف يمكنك استخدام طريقة Deleeuw لتقدير معالم النموذج الموزع متأخراً للفترة k السابق؟

21.17 بالعودة إلى تمرين 15.12 . بما أن قيمة d تظهر قدرة ضعيفة على الكشف عن الارتباط الذاتي (من الدرجة الأولى) (لماذا؟)، كيف يمكنك اختبار الارتباط الذاتي في مثل هذه الحالة؟

مسائل : Problems

22.17 اعتبر النموذج التالي:

$$Y_i^* = \alpha + \beta_0 X_t + u_t$$

بحيث إن Y^* = مصاريف العمل المرغوب فيها، أو على المدى البعيد للحصول على معدات ومصنع جديد، X_t = المبايعات و t = الزمن باستخدام نموذج تعديلات المخزون، قدر معالم دالة الطلب على المصاريف الخاصة بمصنع ومعدات جديدة في المدى البعيد والمعطاة في جدول (8.17). كيف يمكنك التأكد من وجود ارتباط تسلسلي في هذه البيانات؟

جدول (8.17) الاستثمار في معدات وأدوات ثابتة في التصنيع Y ومبايعات التصنيع X_2 (ملايين الدولارات) معدلة موسمياً ، الولايات المتحدة ، 1970 – 1991

Year	Plant expenditure, Y	Sales, X_2	Year	Plant expenditure, Y	Sales, X_2
1970	36.99	52.805	1981	128.68	168.129
1971	33.60	55.906	1982	123.97	163.351
1972	35.42	63.027	1983	117.35	172.547
1973	42.35	72.931	1984	139.61	190.682
1974	52.48	84.790	1985	152.88	194.538
1975	53.66	86.589	1986	137.95	194.657
1976	58.53	98.797	1987	141.06	206.326
1977	67.48	113.201	1988	163.45	223.541
1978	78.13	126.905	1989	183.80	232.724
1979	95.13	143.936	1990	192.61	239.459
1980	112.60	154.391	1991	182.81	235.142

المصدر : Economic Report of the president, 1993.

البيانات عن y من جدول (B-52)، صفحة 407 وبيانات X_2 من جدول صفحة 408

23.17 باستخدام بيانات تمرين 22.17 ولكن باعتبار النموذج التالي :

$$Y_i^* = \beta_0 X_t^{\beta_1} e^{u_t}$$

باستخدام نموذج تعديلات المخزون (لماذا؟)، قدر مرونة الطلب على المعدات الجديدة في المدى البعيد والقصير بالنسبة للمبيعات. قارن نتائجك مع تلك التي حصلت عليها في تمرين 22.17. أي النموذجين تختار ولماذا؟ هل هناك ارتباط تسلسلي في البيانات؟ كيف لك أن تعرف ذلك؟

24.17 استخدم بيانات تمرين 22.17 ولكن افترض أن :

$$Y_t = \alpha + \beta X_t^* + u_t$$

بحيث إن X_t^* هي المبايعات المرغوب فيها. قدر معالم هذا النموذج، وقارن هذه النتائج مع تلك التي حصلت عليها في تمرين 22.17. كيف يمكنك اختيار النموذج المناسب؟ على أساس إحصاء h، كيف يمكنك استنتاج أن هناك ارتباطًا تسلسليًا في هذه البيانات؟

25.17 افترض أن شخصًا ما أقنعك بأن العلاقة بين مصاريف العمل للمعدات الجديدة والمبايعات هي كالتالي :

$$Y_t^* = \alpha + \beta X_t^* + u_t$$

بحيث إن : Y^* هي المصاريف المرغوب فيها و X^* المبايعات المرغوب فيها، استخدم البيانات المعطاة في تمرين 22.17 لتقدير هذا النموذج وعلق على النتائج.

26.17 باستخدام البيانات المعطاة في تمرين 22.17، حدد ما إذا كانت مصاريف المشروع تسبب (Granger) المبايعات أو أن المبايعات تسبب (Granger) مصاريف المشروع. استخدم مالايزيد عن ست فترات زمنية متأخرة وعلق على نتائجك. ما هو الاستنتاج المهم الذي حصلت عليه من هذا التمرين؟

27.17 افترض أن المبايعات في تمرين 22.17 لها تأثير موزع متأخرًا على مصاريف المشروع والمعدات. وفق نموذج Almon للفترات الزمنية المتأخرة لهذه البيانات.

28.17 أعد تقدير المعادلة (16.13.17) بفرض (1) قيد النهاية القريبة (2) قيد النهاية البعيدة و(3) كل من قيدين النهاية وقارن نتائجك المعطاة في (16.13.17). ما الاستنتاج العام الذي خلصت به؟

29.17 جدول 9.17 يعطي بيانات عن استثمار ثابت خاص في عملية المعلومات والمعدات (Y، معطاة بملايين الدولارات)، مبايعات في اجمالي التصنيع

والتجارة (X_2، معطاة بملايين الدولارات) ومعدل فائدة (X_3، معدل السند معطى بـ Moody's Triple A، كنسبة)، بيانات Y، X_2 معدلة موسمياً.

(a) اختبر السببية الثنائية (المزدوجة) بين Y و X_2 مع وضع في الاعتبار طول الفترة الزمنية المتأخرة.

(b) اختبر السببية الثنائية بين Y و X_3 مرة أخرى، مع وضع في الاعتبار طول الفترة الزمنية المتأخرة.

(c) للسماح بظهور أثر الفترات الزمنية المتأخرة للمبايعات على الاستثمار، افترض أنك قررت استخدام طريقة Almon للفترات الزمنية المتأخرة. أوجد النموذج المقدر بعد وضع في الاعتبار طول الفترات الزمنية المتأخرة ودرجة متعددة الحدود المستخدمة.

30.17 جدول (10.17) يعطي بيانات عن مؤشرات التعويض الحقيقي لكل ساعة (Y) والنتائج لكل ساعة (X_2) وكل من المؤشرين للأساسي 100=1992 في قطاع الأعمال في الاقتصاد الأمريكي عن الفترة 1960–1999 بالإضافة إلى معدل البطالة المدنية (X_3) في نفس الفترة.

جدول (9.17) الاستثمار، المبايعات ومعدل الفائدة، الولايات المتحدة 1960–1999

Observation	Investment	Sales	Interest	Observation	Investment	Sales	Interest
1960	4.9	60,827	4.41	1980	69.6	327,233	11.94
1961	5.2	61,159	4.35	1981	82.4	355,822	14.17
1962	5.7	65,662	4.33	1982	88.9	347,625	13.79
1963	6.5	68,995	4.26	1983	100.8	369,286	12.04
1964	7.3	73,682	4.40	1984	121.7	410,124	12.71
1965	8.5	80,283	4.49	1985	130.8	422,583	11.37
1966	10.6	87,187	5.13	1986	137.6	430,419	9.02
1967	11.2	90,820	5.51	1987	141.9	457,735	9.38
1968	11.9	96,685	6.18	1988	155.9	497,157	9.71
1969	14.6	105,690	7.03	1989	173.0	527,039	9.26
1970	16.7	108,221	8.04	1990	176.1	545,909	9.32
1971	17.3	116,895	7.39	1991	181.4	542,815	8.77
1972	19.3	131,081	7.21	1992	197.5	567,176	8.14
1973	23.0	153,677	7.44	1993	215.0	595,628	7.22
1974	26.8	177,912	8.57	1994	233.7	639,163	7.96
1975	28.2	182,198	8.83	1995	262.0	684,982	7.59
1976	32.4	204,150	8.43	1996	287.3	718,113	7.37
1977	38.6	229,513	8.02	1997	325.2	753,445	7.26
1978	48.3	260,320	8.73	1998	367.4	779,413	6.53
1979	58.6	297,701	9.63	1999	433.0	833,079	7.04

لاحظ أن: الاستثمار: استثمار خاص ثابت في عملية المعدات والبرامج، بلايين الدولارات، معدلة موسمياً.

المبايعات: مبايعات في إجمالي التصنيع والتجارة، ملايين الدولارات، معدلة موسمياً.

الفائدة: معدل السند Moody's Triple A، %

المصدر: Economic Report of the president, 2001, tables 13-18, p57 and p.73

جدول (10.17) التبديل ، الإنتاجية ومعدل البطالة ، الولايات المتحدة ، 1960–1999

Observation	COMP	PRODUCT	UNRate	Observation	COMP	PRODUCT	UNRate
1960	60.0	48.8	5.5	1980	89.5	80.4	7.1
1961	61.8	50.6	6.7	1981	89.5	82.0	7.6
1962	63.9	52.9	5.5	1982	90.9	81.7	9.7
1963	65.4	55.0	5.7	1983	91.0	84.6	9.6
1964	67.9	57.5	5.2	1984	91.3	87.0	7.5
1965	69.4	59.6	4.5	1985	92.7	88.7	7.2
1966	71.9	62.0	3.8	1986	95.8	91.4	7.0
1967	73.8	63.4	3.8	1987	96.3	91.9	6.2
1968	76.3	65.4	3.6	1988	97.3	93.0	5.5
1969	77.4	65.7	3.5	1989	95.9	93.9	5.3
1970	78.9	67.0	4.9	1990	96.5	95.2	5.6
1971	80.4	69.9	5.9	1991	97.5	96.3	6.8
1972	82.7	72.2	5.6	1992	100.0	100.0	7.5
1973	84.5	74.5	4.9	1993	99.9	100.5	6.9
1974	83.5	73.2	5.6	1994	99.7	101.9	6.1
1975	84.4	75.8	8.5	1995	99.3	102.6	5.6
1976	86.8	78.5	7.7	1996	99.7	105.4	5.4
1977	87.9	79.8	7.1	1997	100.4	107.6	4.9
1978	89.5	80.7	6.1	1998	104.3	110.5	4.5
1979	89.7	80.7	5.8	1999	107.3	114.0	4.2

لاحظ أن : COMP= مؤشر التبديل الحقيقي لكل ساعة (1992=100)
Product= مؤشر الإنتاج لكل ساعة (1992=100)
UN Rate= معدل البطالة المدنية و %

المصدر : Economic report of the president, 2001, Table B-49, p.332.

(a) كيف يمكنك أن تعرف ما إذا كان بديل الأجر هو الذي يحدد إنتاجية العمالة أو العكس هو الصحيح ؟

(b) أوجد نموذجًا مناسبًا لاختبار تخمينك في a، اعط الإحصاءات المناسبة .

(c) هل تعتقد أن معدل البطالة له أي تأثير على تبديل الأجر ، وإذا كان ذلك صحيحًا كيف يمكنك أن تضع ذلك في الاعتبار عند الدراسة؟ وضح التحليل الإحصائي المطلوب :

31.17 اختبار السببية لـ Sims(*)

في سببية Granger الثنائية ، Sims أوضح حقيقة أن المستقبل لايمكن أن يسبب الماضي . افترض أننا نريد أن نعرف ما إذا كانت X سببًا لـ Y . الآن دعنا نعتبر النموذج التالي :

(*) C.A.Sims, "Money, Income, and Causality," American Economic Review, vol. 62, 1972, pp. 540–552.

$$Y_t = \alpha + \beta_k X_{t-k} + \beta_{k-1} X_{t-k-1} + \cdots + \beta_1 X_{t-1} + \beta_0 X_t$$
$$+ \lambda_1 X_{t+1} + \lambda_2 X_{t+2} + \cdots + \lambda_m X_{t+m} + u_t$$

هذا الانحدار يشتمل على الفترات الزمنية الحالية، المتأخرة والمستقبلية للمتغير المنحدر X، مقادير مثل X_{t+1} و X_{t+2} تسمى مقادير قائدة. في الانحدار السابق، هناك عدد k من الفترات الزمنية المتأخرة، وعدد m من الفترات الزمنية القائدة. إذا كانت X (Granger) تسبب Y فإن مجموع معاملات القيم القائدة لـ X لابد أن تكون إحصائيًا مساوية للصفر(*).

طبق اختبار Sims للسببية على البيانات المعطاة في تمرين 22.17 لتحديد ما إذا كانت المبايعات (Granger) تسبب مصاريف الاستثمار أم لا. حدد بنفسك العدد المناسب من الفترات الزمنية المتأخرة والفائدة للمتغير المنحدر.

APPENDIX ملحق A17

1.A17 اختبار SARGAN لصلاحية المتغيرات المساهمة :
THE SARGAN TEST FOR THE VALIDITY OF INSTRUMENTS

افترض أننا نريد استخدام متغيرات مساهمة بدلاً من متغيرات مفسرة تكون مرتبطة مع مقدار الخطأ. كيف يمكن الحكم على صلاحية هذه المتغيرات المساهمة، بمعنى آخر كيف لنا أن نعرف أن المتغيرات المساهمة المختارة مستقلة عن مقدار الخطأ؟ Sargan اقترح وطور إحصاء ما يرمز له بالرمز SARG لاختبار صلاحية المساهمة في المتغيرات المساهمة (IV)(†). الخطوات الخاصة لـ SARG كالتالي(‡):

(*) الاختيار بين اختبار السببية لـ Sims أو Granger ليس واضحًا. لتفاصيل أكثر عن هذين الاختبارين انظر G.Chamberlain, "The General Equivalence of Granger and Sims Causality." Econometrica, vol 50, 1982, pp.569- 582.

(†) Sargan, J.D., "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in P. E. Hart, G. Mills, and J. K. Whitaker (eds.) Econometic Analysis for National Economic Planning, Butterworths, London, 1964.

(‡) المناقشة التالية مختارة من : H.R.Selddighi, K.A.Lawler and A. V. Katos, Econometrics: A Practical Approach, Routledge, New York, 2000, pp.155-156.

1- قم بتقسيم المتغيرات الموجودة في معادلة الانحدار إلى مجموعتين، مجموعة مستقلة عن مقدار الخطأ (مثلاً X_1، X_2، . . . ، X_p) ومجموعة غير مستقلة عن مقدار الخطأ (مثلاً Z_1، Z_2، ...، Z_q).

2- اجعل W_1، W_2، . . . ، W_5 عبارة عن المتغيرات المساهمة المختارة عن المتغيرات Z الموجودة في 1 بحيث إن $s \geq q$.

3- قدر الانحدار الأصلي مستبدلاً Z's بـ W's، بمعنى آخر قدر الانحدار الأصلي بالـ IV واحصل على البواقي، مثلاً $\hat{u}$.

4- قم بانحدار $\hat{u}$ على ثابت، كل متغيرات الـ X، وكل متغيرات الـ W ولكن استثني المتغيرات Z. احصل على R^2 الخاصة بهذا الانحدار.

5- والآن احسب إحصاء SARG والمعروف كالتالي:

$$\text{SARG} = (n-k)R^2 \qquad \textbf{(1.1.A17)}$$

بحيث إن:

n = عدد المشاهدات، و k = عدد المعاملات في معادلة الانحدار الأصلية. Sargent وضح أن (1.1.A17) يتبع توزيع X^2 بدرجات حرية r بحيث إن $r = s - q$.

6- الفرض العدمي هو أن كل المتغيرات المساهمة w صالحة. إذا كانت قيمة X^2 المحسوبة تزيد عن قيمة λ^2 الحرجة، نرفض الفرض العدمي، بمعنى أن هناك على الأقل متغيراً مساهماً واحداً مرتبطاً مع مقدار الخطأ، وبالتالي فإن تقديرات IV المعتمدة على المتغيرات المساهمة المختارة غير صالحة.

الجزء الرابع

نماذج المعادلات الآنية

SIMULTANEOUS- EQUATION MODELS

النظرة العامة على ما تم نشره من أبحاث تطبيقية في مجال الأعمال والاقتصاد توضح أن العديد من العلاقات الاقتصادية هي من نوع المعادلات الفردية، ولهذا السبب خصصنا الأجزاء الثلاثة الأولى من هذا الكتاب، لدراسة نماذج الانحدار الخاصة بالمعادلة الفردية.

في مثل هذا الإطار، متغير واحد (المتغير التابع Y) يتم التعبير عنه كدالة خطية في متغير أو أكثر في (المتغيرات المفسرة، $X's$). في مثل هذه النماذج، وكفرض ضمني، فإنه إذا وجدت علاقة بين X و Y فإنها تكون علاقة سببية وتتجه في اتجاه واحد:

المتغيرات المفسرة هي السبب، والمتغير التابع هو النتيجة. عمومًا هناك مواقف أخرى يكون فيها تيار التأثير ذا اتجاهين بين المتغيرات الاقتصادية، بمعنى أن متغيرًا اقتصاديًا واحدًا يؤثر على متغير أو متغيرات اقتصادية أخرى، وفي المقابل يتأثر بها.

وبالتالي في انحدار المال M على معدل الفائدة r طريقة المعادلة الفردية تفترض ضمنيًا أن معدل الفائدة ثابت (مثلاً عن طريق نظام الادخار الفيدرالي)، ونحاول إيجاد استجابة الطلب على المال إلى التغير في مستوى معدل الفائدة. ولكن ماذا سيحدث إذا كان معدل الفائدة يعتمد على الطلب على المال؟ في مثل هذه الحالة، فإن تحليل الانحدار الشرطي الذي تم

استعراضه في هذا الكتاب حتى الآن، لايكون مناسبا، حيث إن M تعتمد على r و r تعتمد على M.

وبالتالي، نحن نحتاج على اعتبار معادلتين، واحدة تخلق علاقة لـ M مع r، وأخرى تخلق علاقة لـ r مع M. وهذا يؤدي بنا إلى دراسة نماذج المعادلات الآنية، النماذج التي يكون فيها أكثر من معادلة انحدار واحدة، فيكون هناك معادلة لكل متغير متضامن.

في الجزء الرابع، سنستعرض مقدمة أساسية لموضوع أكثر تعقيداً وهو نماذج المعادلات الآنية، أما التفاصيل فموجودة في المراجع المذكورة. في الفصل (18)، سنستعرض العديد من الأمثلة الخاصة بنماذج المعادلات الآنية، ونوضح لماذا تعتبر طريقة المربعات الصغرى التقليدية بوجه عام غير قابلة للإستخدام في تقدير المعلمات الخاصة بكل معادلة في مثل هذه النماذج.

في الفصل (18)، سنستعرض ما يسمى بمشكلة التوصيف (كشف الهوية)، فإذا كان هناك نظام من المعادلات الآنية يشمل اثنتين أو أكثر من المعادلات سيكون من غير الممكن إيجاد قيم رقمية لكل معامل في كل معادلة، حيث إن المعادلات لايمكن التفريق بينها بشكل ملاحظ أو تشبه المعادلات بعضها البعض إلى حد كبير، وبالتالي تظهر مشكلة التوصيف.

وبالتالي في انحدار الكمية Q على السعر P، هل المعادلة الناتجة دالة طلب أم دالة عرض إذا كان كل من Q و P في كل من المعادلتين؟

وبالتالي، إذا كانت لدينا بيانات عن Q و P فقط، ولا توجد أي معلومات أخرى، ستكون هناك صعوبة أو استحالة للتعرف على الانحدار، وتحديد ما إذا كان معادلة طلب أو عرض.

إنه من الضروري حل مشكلة التعرف قبل التعرض لمسألة التقدير، حيث إذا لم تستطع معرفة ما هو المطلوب تقديره، فإن التقدير نفسه سيكون مضللاً. في الفصل (19) سنعرض عدداً من الطرق المختلفة لحل مشكلة التعرف.

في الفصل (20)، سنستعرض عدداً من طرق التقدير التي صممت خصيصاً لتقدير نماذج المعادلات الآنية، وسندرس حسناتها وقيودها.

الفصل الثامن عشر

نماذج المعادلات الآنية

SIMULTANEOUS- EQUATION MODELS

في هذا الفصل والفصلين التاليين، سنستعرض نماذج المعادلات الآنية، وبالتحديد سندرس الصفات الخاصة بها، تقديرها، وبعض المشاكل الإحصائية المرتبطة بها.

1.18 طبيعة نماذج المعادلات الآنية :
THE NATURAL OF SIMULTANEOUS- EQUATION MODELS

أجزاء الكتاب من جزء I إلى جزء III، خاصة بنماذج المعادلة الفردية، بمعنى النماذج، التي يكون فيها متغير تابع واحد Y واحد أو أكثر من المتغيرات المفسرة، X's.

في مثل هذه النماذج ، فإن الهدف كان تقدير أو التنبؤ بمتوسط قيمة y بشرط قيم محددة للمتغيرات المفسرة X. علاقة السبب والنتيجة إن وجدت في مثل هذه النماذج، تكون من الـ X's في اتجاه الـ Y.

ولكن في العديد من المواقف، مثل هذه العلاقة الأحادية الاتجاه للسبب والنتيجة تكون بدون معنى. وهذا يحدث إذا كانت Y تحدد بالـ X's وبعض الـ X's في المقابل تحدد بالـ Y.

باختصار، فإن هناك اتجاهين أو علاقة آنية بين Y و(بعض) الـ X's، والذي يجعل هناك صعوبة في التفرقة بين المتغيرات التابعة والمستقلة، مما يؤدي إلى التباس في القيم الخاصة بها. من الأفضل إجمال مجموعة من المتغيرات التي يمكن أن تحدد في نفس الوقت معًا عن طريق باقي المتغيرات، وهذا ما يتم بدقة في نماذج المعادلات

الآنية. في مثل هذه النماذج، هناك أكثر من معادلة واحدة لكل المتغيرات المدمجة التابعة أو الداخلية الموجودة في النموذج[1]. وبخلاف نماذج المعادلة الفردية، فإن نماذج المعادلات الآنية قد لا تستطيع تقدير معلمات معادلة واحدة، بدون أن يؤخذ في الاعتبار معلومات مأخوذة من معادلات أخرى داخل النظام.

ماذا سيحدث إذا تم تقدير معاملات كل معادلة بتطبيق مثلاً طريقة الـ OLS بغض النظر عن المعادلات الأخرى الموجودة في النظام؟ تذكر أن أحد الافتراضات المهمة في طريقة الـ OLS، أن المتغيرات المفسرة إما ثابتة أو إذا كانت عشوائية فإنها تكون موزعة بشكل مستقل عن مقدار الخطأ (التشتت) العشوائي. إذا لم يتم تحقق أي من هذه الشروط، فإنه كما سبق وأوضحنا، تقديرات المربعات الصغرى ليست فقط متحيزة، وإنما أيضًا غير متسقة، بمعنى أنه كلما يزداد حجم العينة فإن التقديرات لاتؤول إلى قيمها الحقيقية (قيم المجتمع). وبالتالي في نظام المعادلات الافتراضي التالي[2]:

$$Y_{1i} = \beta_{10} + \beta_{12}Y_{2i} + \gamma_{11}X_{1i} + u_{1i} \qquad (1.1.18)$$

$$Y_{2i} = \beta_{20} + \beta_{21}Y_{1i} + \gamma_{21}X_{1i} + u_{2i} \qquad (2.1.18)$$

بحيث إن Y_1 و Y_2 مستقلان تبادليًا أو متغيرات داخلية، و X_1 متغير خارجي، و u_1 و u_2 مقادير الأخطاء العشوائية، المتغيرات Y_1، Y_2 كلاهما متغيران عشوائيان.

وبالتالي إذا لم تستطع إيضاح أن المتغير المفسر العشوائي Y_2 في (1.1.18) موزعًا مستقلاً عن u_1، والمتغير المفسر العشوائي Y_1 في (2.1.18) موزع مستقلاً عن u_2 فإن تطبيق OLS التقليدية لمثل هذه المعادلات بشكل منفرد سيؤدي إلى مقدرات غير مستقلة.

في باقي هذا الفصل، سنعطي بعض الأمثلة لنماذج المعادلات الآنية، ونوضح التحيز الذي سيظهر عند التطبيق المباشر لطريقة المربعات الصغرى لمثل هذه النماذج. وبعد مناقشة مايسمى بمشكلة التعرف في الفصل (19). في الفصل (20) سنستعرض بعض النماذج التي صممت للتعامل مع نماذج المعادلات الآنية.

---

(1) في موضوع نماذج المعادلات الآنية، فإن المتغيرات التابعة تسمى متغيرات داخلية، والمتغيرات التي تعتبر ساكنة بشكل حقيقي أو يمكن اعتبارها ساكنة تسمى متغيرات خارجية أو سابقة التحديد (المزيد من التفاصيل في الفصل 19).

(2) هذا الترميز الاقتصادي والمفسر ذاتيًا سيتم تصميمه عند استخدام أكثر من معادلتين في الفصل 19.

2.18 أمثلة لنماذج المعادلات الآنية :
EXAMPLES OF SIMULTANEOUS- EQUATION MODELS

مثال 1.18

نموذج العرض والطلب : **Example 18.1 Demand-and- Supply Model**

كما هو معروف، فإن السعر P لسلعة ما، والكمية المباعة Q يتم تحديدهما عند نقطة التقاطع بين منحنيات العرض والطلب الخاصة بهذه السلعة. وبالتالي دعنا نفترض للتبسيط أن منحنيات العرض والطلب خطية، ونصف مقادير الأخطاء u_1 و u_2، وبالتالي يمكن كتابة معادلات العرض والطلب كالتالي:

(1.2.18) دالة العرض: $\alpha_1 < 0$ $\quad Q_t^d = \alpha_0 + \alpha_1 P_t + u_{1t}$

(2.2.18) دالة الطلب: $\beta_1 > 0$ $\quad Q_t^s = \beta_0 + \beta_1 P_t + u_{2t}$

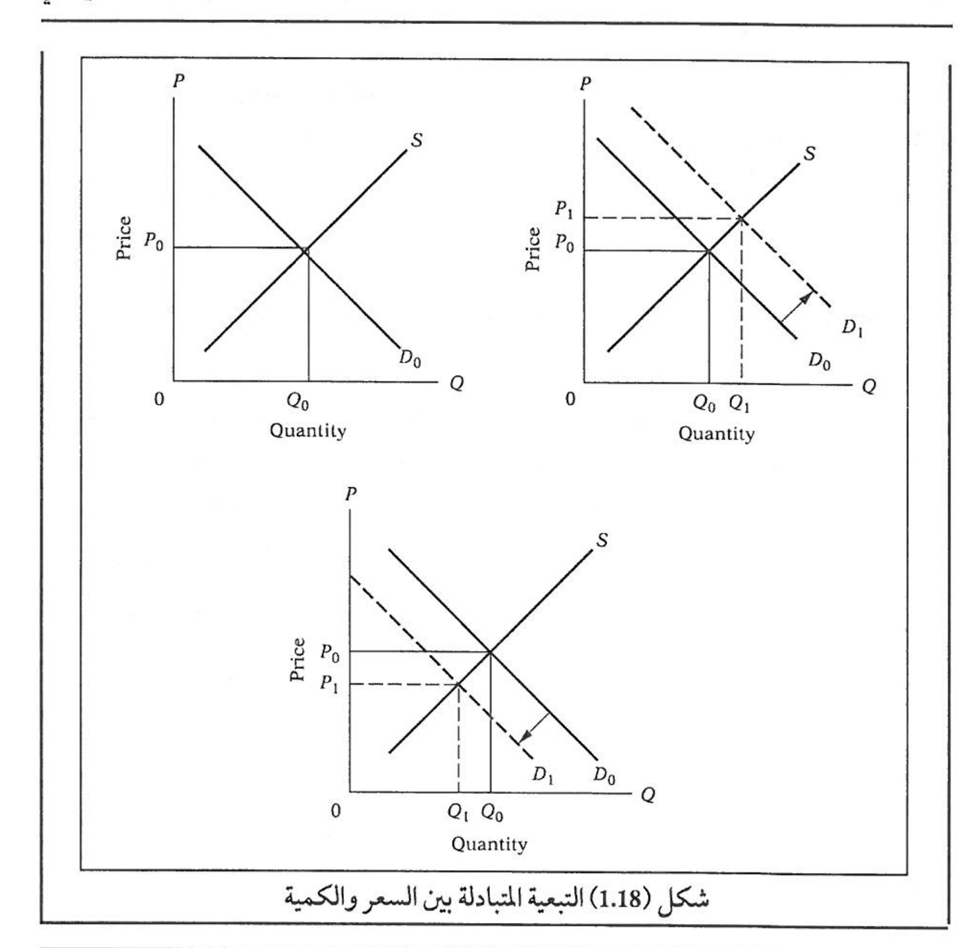
شرط التوازن: $Q_t^d = Q_t^s$

بحيث إن: Q^d = الكمية المطلوبة
Q^s = الكمية المعروضة
t = الزمن

والـ α's والـ β's هي المعاملات. مسبقًا فإن α_1 متوقع أن تكون سالبة (الاتجاه العكسي لمنحنى الطلب)، و β_1 متوقع أن يكون موجبًا (الاتجاه الطردي لمنحنى العرض).

الآن لن يكون صعبًا أن ترى أن P و Q متغيرات تابعة مشتركة. فمثلاً إذا تغير u_{1t} في (1.2.18) بسبب التغير في المتغيرات الأخرى التي تؤثر على Q_t^d (مثل الدخل، الثروة، والمذاق) فإن منحنى الطلب سينتقل إلى أعلى، إذا كان u_{1t} موجبًا وإلى أسفل، إذا كان u_{1t} سالبًا هذه الانتقالات موضحة في شكل (1.18).

وكما يوضح الشكل، فإن الانتقال في منحنى الطلب يغير كلاً من P و Q. وبالمثل فإن التغير في u_{2t} (بسبب العواصف، المناخ، قيود الاستيراد والتصدير وهكذا) سينقل منحنى العرض، مرة أخرى مؤثرًا على P و Q. وبسبب هذه التبعية الآنية بين Q و P، u_{1t} و P_t في (1.2.18) و u_{2t} و P_t في (2.2.18) لايمكن أن يكونا مستقلين. وبالتالي انحدار Q على P كما في (1.2.18) سيخالف فرضًا مهمًا في نموذج الانحدار الخطي التقليدي، وهو فرض عدم وجود ارتباط بين المتغيرات المفسرة ومقدار الخطأ.



شكل (1.18) التبعية المتبادلة بين السعر والكمية

مثال 2.18

نموذج Keynesian لتحديد الدخل :

Keynesian Model of Income Determination

اعتبر النموذج البسيط لـ Keynesian لتحديد الدخل التالي :

(3.2.18) دالة الاستهلاك $\quad C_t = \beta_0 + \beta_1 Y_t + u_t \qquad 0< \beta_t < 1$

(4.2.18) وحدة الدخل $\quad Y_t = C_t + I_t \, (= S_t)$

بحيث إن :

C = مصاريف الاستهلاك

Y = الدخل

I= الاستثمار (مفترض كمتغير خارجي)

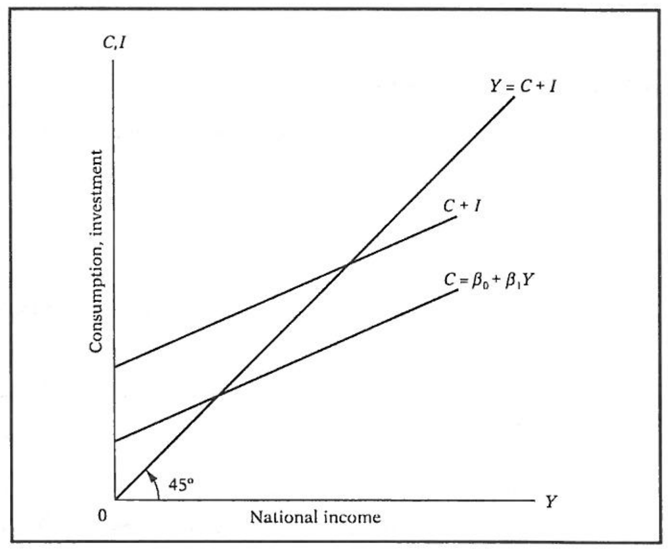
S= الادخار

T= الزمن

U= مقدار التشتت

β_0 و β_1= معاملات الانحدار

المعامل β_1 معروف على أنه الميل الحدي للاستهلاك (MPC) (الكمية الزائدة في مصاريف الاستهلاك والناتجة عن زيادة دولار واحد في الدخل). من النظرية الاقتصادية β_1 مفترض أن تكون بين 0 و 1. المعادلة (3.2.18) هي دالة الاستهلاك (العشوائية) و(4.2.18) هي وحدة الدخل القومي، بمعلومية أن إجمالي الدخل يساوي مصاريف الاستهلاك الكلية مضافًا إليها مصاريف الاستثمار، ومن المفهوم أن مصاريف الاستثمار تساوي إجمالي الدخل. شكل (2.18) يعتبر جدولاً بيانيًا يعبر عن كل ذلك. من دالة الاستهلاك المفترضة وشكل (2.18) من الواضح أن C و Y مستقلان تبادليًا و Y_t الموجودة في (3.2.18) من غير المتوقع أن تكون مستقلة عن مقدار التشتت (الخطأ)، حيث عندما ينتقل u_t (بسبب العديد من العوامل المؤثرة المتضمنة في مقدار الخطأ) فإن دالة الاستهلاك تنتقل هي الأخرى، والذي في المقابل يؤثر على Y_t. بالتالي فإنه مرة أخرى يكون من غير الممكن تطبيق طريقة المربعات الصغرى التقليدية للنموذج (3.2.18) وإذا تم تطبيقها، فإن المقدرات التي سيتم الحصول عليها ستكون غير متسقة كما سنرى لاحقًا.



شكل (2.18) نموذج Keynesian لتحديد الدخل

مثال 3.18

نموذج الأجر– السعر : Wage- Price Model

اعتبر النموذج التالي من نوع Phillips والخاص بتحديد ثمن الأجر والسعر.

$$\dot{W}_t = \alpha_0 + \alpha_1 \text{UN}_t + \alpha_2 \dot{P}_t + u_{1t} \qquad (5.2.18)$$

$$\dot{P}_t = \beta_0 + \beta_1 \dot{W}_t + \beta_2 \dot{R}_t + \beta_3 \dot{M}_t + u_{2t} \qquad (6.2.18)$$

بحيث إن: $\dot{W}$ = معدل التغير في الأجر

UN = معدل البطالة، %

$\dot{P}$ = معدل التغير في الأسعار

$\dot{R}$ = معدل التغير في تكلفة رأس المال

$\dot{M}$= معدل التغير في سعر المواد الخام المستوردة

t = الزمن

u_1، u_2= مقادير التشتت (الأخطاء) العشوائية

بما أن سعر المتغير $\dot{P}$ يدخل في معادلة الأجر، وفي نفس الوقت متغير الأجر $\dot{W}$ يدخل في معادلة السعر، فإن المتغيرين هما متغيران تابعان بالتبادل. وبالتالي، فإن المتغيرات المفسرة المتغيرة الخاصة بها من المتوقع أن يكونا مرتبطين مع مقادير التشتت العشوائية الخاصة بهما، ومرة أخرى فإن طريقة OLS التقليدية لايمكن تطبيقها لتقدير المعادلات الخاصة بتلك المعادلتين منفردتين.

مثال 4.18

نموذج IS للاقتصاد الكلي : The IS Model of Macroeconomics

نموذج الاقتصاد الكلي المعروف IS أو توازن سلعة السوق(3) يمكن التعبير عنه كالتالي (بشرط وجوده في حالة غير العشوائية):

(7.2.18) دالة استهلاك	$0 < \beta_1 < 1$	$C_t = \beta_0 + \beta_1 Y_{dt}$
(8.2.18) دالة الضرائب	$0 < \alpha_1 < 1$	$T_t = \alpha_0 + \alpha_1 Y_t$
(9.2.18) دالة استثمار		$I_t = \gamma_0 + \gamma_1 r_t$
(10.2.18) التعريف		$Y_{dt} = Y_t - T_t$
(11.2.18) مصاريف حكومية		$G_t = \bar{G}$
(12.2.18) وحدة الدخل القومي		$Y_t = C_t + I_t + G_t$

(3) "جدول توازن سلع السوق أو جدول IS يوضح التوليفات المختلفة بين معدلات الفائدة ومستويات الإنتاج، بحيث يكون الإنفاق يساوي الدخل". انظر:

Rudiger Dornbusch and Stanley Ficher, Macroeconomic, 3d. ed., McGraw-Hill, New York, 1984, p.102

لاحظ أنه للتبسيط افترضنا استبعاد قطاع التجارة الخارجية.

بحيث إن : Y = الدخل القومي
C = الصرف على الاستهلاك
I = الاستثمار الصافي المخطط له أو المرغوب فيه
$\overline{G}$= مستوى محدد للمصاريف الحكومية
T = الضرائب
Y_d= الدخل المتاح
r = معدل الفائدة

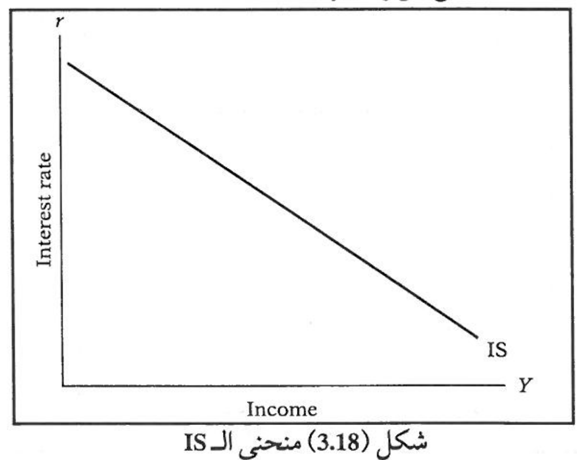
إذا عوضنا بـ (10.2.18) و (8.2.18) في (7.2.18) وعوضنا بالمعادلة الناتجة لـ C ومعادلة (9.2.18) و (11.2.18) في (12.2.18) سنحصل على :

(13.2.18) معادلة IS : $$Y_t = \pi_0 + \pi_1 r_t$$

بحيث إن : $$\pi_0 = \frac{\beta_0 - \alpha_0\beta_1 + \gamma_0 + \bar{G}}{1 - \beta_1(1 - \alpha_1)}$$

(14.2.18) $$\pi_1 = \frac{1}{1 - \beta_1(1 - \alpha_1)}$$

معادلة (13.2.18) هي معادلة IS، أو توازن سلع السوق، بمعنى أنها تعطي التوليفة المطلوبة بين معدل الفائدة ومستوى الدخل ، بحيث إن سلع السوق تتوازن هندسيًا. منحنى IS موضح في الشكل (3.18) ماذا سيحدث إذا كنا نقدر مثلاً دالة الاستهلاك (7.2.18) منفردة؟ هل من الممكن أن نحصل على مقدر غير متحيز أو متسق أو كليهما معًا لكل من β_0 و β_1؟ هذه النتيجة غير محتملة ، حيث إن الاستهلاك يعتمد على الدخل المتاح ، والذي بدوره يعتمد على الدخل القومي Y ولكن الأخير يعتمد على r و $\overline{G}$ بالإضافة إلى معاملات أخرى تدخل في IIO ، وبالتالي إذا لم تأخذ في الاعتبار كل هذا الأثر ، فإن الانحدار البسيط لـ C على Y_d سيعطي تقديرات متحيزة أو غير متسقة أو كليهما معًا لكل من β_0 و β_1 .



شكل (3.18) منحنى الـ IS

مثال 5.18

نموذج الـ LM : The LM model

النصف الآخر المشهور لنموذج IS-LM هو LM أو علاقة توازن المال في السوق، والذي يعطي التوليفة المطلوبة بين معدل الفائدة ومستوى الدخل، بحيث إن المال في السوق يتوازن، بمعنى أن الطلب على المال يساوي المعروض عليه. جبريًا فإن النموذج في صورته غير العشوائية يمكن التعبير عنه كالتالي:

(15.2.18) دالة الطلب على المال $M_t^d = a + bY_t - cr_t$

(16.2.18) دالة العرض على المال $M_t^s = \bar{M}$

(17.2.18) شرط التوازن $M_t^d = M_t^s$

بحيث إن Y = الدخل، r = معدل الفائدة و$\overline{M}$ = المستوى المفترض للمعروض من المال مثلاً المحدد فيدراليًا.

تساوي دالة الطلب والعرض على المال وتبسيطها نحصل على:

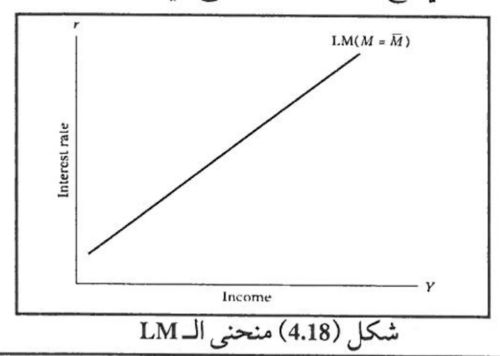
(18.2.18) معادلة LM: $Y_t = \lambda_0 + \lambda_1 \bar{M} + \lambda_2 r_t$

بحيث إن:

$$\lambda_0 = -a/b$$
$$\lambda_1 = 1/b \qquad (19.2.18)$$
$$\lambda_2 = c/b$$

عند مستوى محدد لـ M مثلاً $\overline{M}$ = M، فإن منحنى LM الذي يوضح العلاقة (18.2.18) معطى في شكل (4.18).

منحنا IS و LM يوضحان بالتوالي متجه كل من معدلات الفائدة متسقًا مع سلع السوق المتوازنة ومتجه كلي من معدلات الفائدة المتلائم مع التوازن في سوق المال. بالطبع قيمة واحدة فقط من معدل الفائدة ومستوى واحد فقط من الدخل سيكون متسقًا بشكل آني مع نقطتي التوازن. وللحصول على ذلك كل ما نحتاج لعمله هو مساواة (13.2.18) مع (18.2.18). في تمرين 4.18 مطلوب توضيح مستوى معدل الفائدة والدخل الملائمين بشكل آني مع توازن المال والسلع في السوق.



شكل (4.18) منحنى الـ LM

مثال 6.18

نماذج الاقتصاد القياسي : Econometric Models

نماذج المعادلات الآنية تم استخدامها بشكل خاص في نماذج الاقتصاد القياسي على يد العديد من علماء الاقتصاد القياسي. من الأوائل في هذا المجال العالم Lawrence Klein من مدرسة Wharton في جامعة بنسلفانيا University of Pennsylivania . نموذجه المبدئي معروف باسم نموذج klein والموجود على شكل:

دالة الاستهلاك: $C_t = \beta_0 + \beta_1 P_t + \beta_2 (W + W')_t + \beta_3 P_{t-1} + u_{1t}$

دالة الاستثمار: $I_t = \beta_4 + \beta_5 P_t + \beta_6 P_{t-1} + \beta_7 K_{t-1} + u_{2t}$

الطلب على العمالة: $W_t = \beta_8 + \beta_9 (Y + T - W')_t + \beta_{10} (Y + T - W')_{t-1} + \beta_{11} t + u_{3t}$ **(20.2.18)**

عرف: $Y_t + T_t = C_t + I_t + G_t$

عرف: $Y_t = W'_t + W_t + P_t$

عرف: $K_t = K_{t-1} + I_t$

بحيث إن: C = مصاريف الاستهلاك

I = مصاريف الاستثمار

G = مصاريف حكومية

P = الربح

W = الأجر الخاص

W' = الأجر الحكومي

K = أسهم رأس المال

T = الضرائب

Y = الدخل بعد الضريبة

t = الزمن

u_1، u_2 و u_3 = مقادير التشتت (الخطأ) العشوائي[4].

في النموذج السابق المتغيرات C، I، W، Y، P و K يتم التعامل معها على أنها تابعة تبادلياً أو متغيرات داخلية، المتغيرات P_{t-1}، K_{t-1}، Y_{t-1} يتم التعامل معها على إنها محددة سلفاً[5].

(4) L.R.Klein, Economic Fluctuations in the United States, 1921–1941, John Wiley & Sons, New York, 1950.

(5) الباحث الذي يبني النموذج عليه تحديد المتغيرات في النموذج والتي تعتبر داخلية والأخرى التي تعتبر محددة سلفاً. K_{t-1} و Y_{t-1} محددين سلفاً حيث إنه عند النقطة الزمنية t فإن قيمها تكون معروفة (المزيد من التفاصيل في الفصل 19).

وكإجمالي، فإن لدينا ست معادلات (بالإضافة إلى التعاريف الثلاثة السابقة) لدراسة علاقة التبعية التبادلية للمتغيرات الستة الداخلية.

في الفصل (20)، سنرى كيف تقدر مثل هذه النماذج في الاقتصاد القياسي، الآن لاحظ أنه بسبب التبعية المتبادلة بين المتغيرات الداخلية فإنها بوجه عام لا تعتبر متغيرات مستقلة عن مقادير الخطأ العشوائية والتي تجعل هناك عدم إمكانية لتطبيق طريقة الـ OLS للمعادلات الموجودة في النظام بشكل منفرد. كما هو موضح في الفقرة 3.18، المقدرات التي يتم الحصول عليها ستكون غير متسقة فلا تؤول إلى قيم المجتمع الحقيقية حتى وان كان حجم العينة كبيراً.

3.18 تحيز المعادلات الآنية.. عدم اتساق مقدرات الـ OLS : THE SIMULTANEOUS EQUATION BIAS: INCONSISTENCY OF OLS ESTIMATORS

كما سبق وذكرنا، فإن طريقة المربعات الصغرى قد لايمكن تطبيقها لتقدير المعادلات الموجودة في نظام المعادلات الآنية. كل منها بشكل منفرد بسبب ارتباط واحد أو أكثر من المتغيرات المفسرة مع مقدار الخطأ الموجود في المعادلة، حيث إن المقدرات التي سيتم الحصول عليها ستكون غير متسقة.

لإيضاح ذلك، دعنا نستعرض نموذج Keynesian المبسط لتحديد الدخل والمعطى في مثال 2.18. افترض أننا نريد أن نقدر معاملات دالة الاستهلاك (3.2.18). بافتراض أن $E(u_t) = 0$، $E(u_t^2) = \sigma^2$، $E(u_t u_{t+j}) = 0$ for $(j \neq 0)$ و $cov(I_t, u_t) = 0$ والتي جميعًا تمثل فروض نموذج الانحدار الخطي التقليدي، أولاً دعنا نوضح أن Y_t و u_t في (3.2.18) مرتبطان ثم نثبت أن $\hat{\beta}_1$ تعتبر مقدراً غير متسق لـ β_1.

لإثبات أن Y_t و u_t مرتبطتان دعنا نقوم بالتالي. عوض بـ (3.2.18) في (4.2.18) فنحصل على:

$$Y_t = \beta_0 + \beta_1 Y_t + u_t + I_t$$

$$Y_t = \frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t + \frac{1}{1-\beta_1} u_t \quad (1.3.18)$$

والآن

$$E(Y_t) = \frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t \quad (2.3.18)$$

حيث يتم استخدام $E(u_t)=0$ و I_t والتي هي متغيرات خارجية أو محددة سلفًا (حيث إنها محددة سلفًا) لها توقع يساوي I_t.

وبالتالي بطرح (2.3.18) من (1.3.18) نحصل على:

$$Y_t - E(Y_t) = \frac{u_t}{1-\beta_1} \qquad (3.3.18)$$

بالإضافة إلى:

$$u_t - E(u_t) = u_t \qquad \text{(لماذا؟)} \qquad (4.3.18)$$

وبالتالي:

$$\text{cov}(Y_t, u_t) = E[Y_t - E(Y_t)][u_t - E(u_t)]$$

$$= \frac{E(u_t^2)}{1-\beta_1} \quad \text{نحصل عليها من (3.3.18) و (4.3.18)}$$

$$= \frac{\sigma^2}{1-\beta_1} \qquad (5.3.18)$$

بما أن σ^2 موجبة افتراضيًا (لماذا؟)، التغاير بين Y و u المعطى في (5.3.18) مفترض أن يأخذ قيمة تختلف عن الصفر(6). وكنتيجة لذلك، فإن Y_t و u_1 في (3.2.18) متوقع أن يكونا مرتبطين، مما يخالف فرض نموذج الانحدار الخطي التقليدي، حيث إن مقادير الأخطاء تكون مستقلة أو على الأقل غير مرتبطة مع المتغيرات المفسرة، كما لاحظنا من قبل، مقدرات OLS في مثل هذه المواقف تكون غير متسقة.

لإثبات أن مقدار الـ OLS لـ$\hat{\beta}_1$ مقدر غير متسق لـ β_1 بسبب الارتباط الموجود بين Y_t و u_t، نقوم بالتالي: -

$$\hat{\beta}_1 = \frac{\sum(C_t - \bar{C})(Y_t - \bar{Y})}{\sum(Y_t - \bar{Y})^2}$$

$$= \frac{\sum c_t y_t}{\sum y_t^2} \qquad (6.3.18)$$

$$= \frac{\sum C_t y_t}{\sum y_t^2}$$

(6) ستكون أكبر من الصفر بما أن β_1، MPC تقع بين 0 و 1 وستكون سالبة إذا كان β_1 أكبر من الوحدة. بالطبع قيمة من قيم MPC تكون أكبر من الوحدة لن يكون له معنى اقتصادي. وبالتالي في الحقيقة فإن التغاير بين Y_t و u_t من المتوقع أن يكون موجبًا.

حيث إن الحروف الصغيرة كالعادة تعبر عن انحراف عن القيمة المتوقعة (العينة) بالتعويض عن C_t من (3.2.18)، نحصل على:

$$\hat{\beta}_1 = \frac{\sum(\beta_0 + \beta_1 Y_t + u_t)y_t}{\sum y_t^2} = \beta_1 + \frac{\sum y_t u_t}{\sum y_t^2} \quad (7.3.18)$$

حيث إنه في الخطوة الأخيرة، تم استخدام حقيقة أن $\sum y_t = 0$ و $(\sum Y_t y_t / \sum y_t^2) = 1$ (لماذا؟).

إذا أخذنا توقع (7.3.18) على كل من الطرفين، نحصل على:

$$E(\hat{\beta}_1) = \beta_1 + E\left[\frac{\sum y_t u_t}{\sum y_t^2}\right] \quad (8.3.18)$$

من سوء الحظ، أننا لانستطيع حساب $E(\sum Y_t u_t / \sum y_t^2)$ بما أن معامل التوقع هو معامل خطي (لاحظ: $E(A/B) \neq E(A)/E(B)$). ونلاحظ بشكل واضح أنه إذا لم يكن المقدار $(\sum Y_t y_t / \sum y_t^2)$ مساويًا للصفر، فإن $\hat{\beta}_1$ مقدر متحيز لـ β_1. ولكن هل أثبتنا بعد في (5.3.18) أن التغاير بين Y و u ليس صفرًا وبالتالي $\hat{\beta}_1$ غير متحيزة؟ الإجابة ليس بالضبط، حيث إن $\text{Cov}(Y_t, u_t)$ بمفهوم المجتمع ليس بالضرورة $\sum y_t u_t$ والذي هو مقياس من العينة، ولكن بزيادة حجم العينة، فإن الأخير سيؤول إلى قيمة المجتمع. وبالتالي نحن نحتاج إلى دراسة الوضع في حالة زيادة حجم العينة هل سيتحقق مفهوم الاتساق ودراسة ماذا سيحدث لـ $\hat{\beta}_1$ مع زيادة حجم العينة.

باختصار، عندما لانستطيع تقييم بوضوح القيمة المتوقعة للمقدار كما في (8.3.18) فإننا يمكن أن نهتم أكثر بسلوك هذا المقدر في العينات كبيرة الحجم.

لأن المقدر يقال عنه إنه متسق إذا كانت النهاية الاحتمالية[7] أو Plim للاختصار تساوي قيمته الحقيقية (المجتمع). وبالتالي لإثبات أن $\hat{\beta}_1$ الموجود في (7.3.18) مقدر غير متسق، فيجب أن نثبت أن Plim الخاصة به لاتساوي القيمة الحقيقية لـ β_1.

ويتطبيق قواعد النهايات الاحتمالية لـ (7.3.18)، نحصل على[8]:

---

(7) انظر APP.A لتعريف النهاية الاحتمالية.

(8) كما هو موجود في APP.A، فإن plim ثابت (على سبيل المثال، $\beta 1$) هي نفس الثابت و plim لـ $(A|B) = \text{plim}(A)/\text{plim}(B)$. لاحظ عمومًا أن $E(A/B) \neq E(A)/E(B)$.

$$\begin{aligned}
\text{plim}\,(\hat{\beta}_1) &= \text{plim}\,(\beta_1) + \text{plim}\left(\frac{\sum y_t u_t}{\sum y_t^2}\right) \\
&= \text{plim}\,(\beta_1) + \text{plim}\left(\frac{\sum y_t u_t/n}{\sum y_t^2/n}\right) \\
&= \beta_1 + \frac{\text{plim}\,(\sum y_t u_t/n)}{\text{plim}\,(\sum y_t^2/n)}
\end{aligned} \tag{9.3.18}$$

حيث إنه في الخطوة الثانية، فقد قسمنا $\sum y_t u_t$ و $\sum y_t^2$ على العدد الكلي للمشاهدات في العينة n، وبالتالي فإن الكميات بين الأقواس تمثل الآن التغاير داخل العينة بين Y و u وتباين العينة لـ Y بالتوالي.

في كلمات (9.3.18) تعني أن النهاية الاحتمالية لـ $\hat{\beta}$ تساوي القيمة الحقيقية لـ β_1 مضافًا إليها النسبة بين plim لتغاير العينة بين Y و u إلى plim لتباين العينة لـ Y. الآن كلما يزداد حجم العينة n، فإنه من المتوقع أن يؤول تغاير العينة بين Y و u على تغاير المجتمع $E[Y_t - E(Y_t)][u_t - E(ut)]$، والذي يساوي $\sigma^2/1-\beta_1$ من (5.3.18). بالمثل كلما تقترب n من المالا نهاية، فإن تباين العينة لـ Y سيقترب من تباين المجتمع (σ_Y^2). وبالتالي المعادلة (9.3.18) من الممكن كتابتها كالتالي:

$$\begin{aligned}
\text{plim}\,(\hat{\beta}_1) &= \beta_1 + \frac{\sigma^2/(1-\beta_1)}{\sigma_Y^2} \\
&= \beta_1 + \frac{1}{1-\beta_1}\left(\frac{\sigma^2}{\sigma_Y^2}\right)
\end{aligned} \tag{10.3.18}$$

علمًا بأن $0 < \beta_1 < 1$ و σ^2 و σ_Y^2 كليهما موجب، من الواضح من المعادلة (10.3.18) أن plim $(\hat{\beta}_1)$ سيكون دائمًا أكبر من β_1، بمعنى أن $\hat{\beta}_1$ ستقدر القيمة الحقيقية لـ β_1 بأكبر من قيمتها الحقيقية[9]. بمعنى آخر، فإن $\hat{\beta}_1$ مقدر متحيز، وهذا التحيز لن يختفي حتى إذا كان حجم العينة كبيراً.

(9) بوجه عام فإن اتجاه التحيز سيعتمد على شكل النموذج والقيم الحقيقية لمعاملات الانحدار.

4.18 تحيز المعادلات الآنية.. مثال رقمي:
THE SIMULTANEOUS EQUATION BIAS: A NUMERICAL EXAMPLE

لتوضيح بعض النقاط التي تم استعراضها في الفقرة السابقة، دعنا نعود إلى نموذج Keynesian البسيط لتحديد الدخل والمعطى في مثال 2.18، ودعنا نقوم بدراسة Monte Carlo التالية(10). افترض أن قيم الاستثمار I معطاة في جدول (1.18).

جدول (1.18)

Y_t (1)	C_t (2)	I_t (3)	u_t (4)
18.15697	16.15697	2.0	−0.3686055
19.59980	17.59980	2.0	−0.8004084E-01
21.93468	19.73468	2.2	0.1869357
21.55145	19.35145	2.2	0.1102906
21.88427	19.48427	2.4	−0.2314535E-01
22.42648	20.02648	2.4	0.8529544E-01
25.40940	22.80940	2.6	0.4818807
22.69523	20.09523	2.6	−0.6095481E-01
24.36465	21.56465	2.8	0.7292983E-01
24.39334	21.59334	2.8	0.7866819E-01
24.09215	21.09215	3.0	−0.1815703
24.87450	21.87450	3.0	−0.2509900E-01
25.31580	22.11580	3.2	−0.1368398
26.30465	23.10465	3.2	0.6092946E-01
25.78235	22.38235	3.4	−0.2435298
26.08018	22.68018	3.4	−0.1839638
27.24440	23.64440	3.6	−0.1511200
28.00963	24.40963	3.6	0.1926739E-02
30.89301	27.09301	3.8	0.3786015
28.98706	25.18706	3.8	−0.2588852E-02

المصدر: Kenneth J.white, Nancy G.., September 1985, p.132

العمود 3 في جدول (1.18) بالإضافة إلى ذلك افترض أن:

$$E(u_t) = 0$$
$$E(u_t u_{t+j}) = 0 \qquad (j \neq 0)$$
$$\text{var}(u_t) = \sigma^2 = 0.04$$
$$\text{cov}(u_t, I_t) = 0$$

الـ u_1 المولدة موجودة في العمود (4).

(10) هذا مأخوذ من: Kenneth J.white, Nancy G. Horsman, and Justin B. Wyatt, SHAZAM: Computer Handbook for Econometrics for Use with Basic Econometrics, McGraw-Hill, New York, 1985, pp. 131-134

بالنسبة لدالة الاستهلاك (3.2.18) افترض أن القيم الحقيقية للمعلمات معروفة وهي $\beta_0=2$ و $\beta_1=0.8$.

من القيم المفترضة لـ β_0 و β_1 والقيم المولدة لـ u_t يمكن أن نولد قيم الدخل Y_t من (1.3.18) والموجودة في العمود 1 في جدول (1.18). بمجرد أن تكون Y_t معلومة، معرفة β_0، β_1 و u_t فمن الممكن توليد قيم الاستهلاك C_t من (3.2.18). الـ C's المولدة موجودة في العمود 2.

بما أن القيم الحقيقية لـ β_0 و β_1 معروفة، وبما أن قيم أخطاء العينة مساوية تمامًا للأخطاء الحقيقية (بسبب الطريقة التي تم تصميم دراسة Monte Carlo)، فإذا استخدمنا بيانات جدول (1.18) لعمل انحدار لـ C_t على Y_t فإنه يجب أن نحصل على $\beta_0=2$، $\beta_1=0.8$ إذا كانت OLS غير متحيزة. ولكن من (7.3.18) نعرف أن ذلك لن يحدث إذا كان المنحدر عليه Y_t ومقدار الخطأ u_t مرتبطين. والآن لن يكون صعبًا إثبات أن تغاير العينة بين Y_t و u_t هو $\Sigma y_t u_t = 3.8$ (من بيانات العينة) وأيضًا $\Sigma y_t^2 = 184$.

وبالتالي كما توضح (7.3.18)، يجب أن يكون لدينا:

$$\hat{\beta}_1 = \beta_1 + \frac{\sum y_t u_t}{\sum y_t^2}$$

$$= 0.8 + \frac{3.8}{184} \qquad \textbf{(1.4.18)}$$

$$= 0.82065$$

بمعنى أن $\hat{\beta}_1$ متحيزة، وتعطي قيمًا أكبر من الحقيقة بحوالي 0.02065، الآن دعنا نقوم بانحدار C_t على Y_t، باستخدام البيانات الموجودة في جدول (1.18). نتائج الانحدار هي كالتالي:

$$\begin{aligned} \hat{C}_t &= 1.4940 \quad + \quad 0.82065 Y_t \\ \text{se} &= (0.35413) \qquad (0.01434) \qquad\qquad \textbf{(2.4.18)} \\ t &= (4.2188) \qquad (57.209) \qquad R^2 = 0.9945 \end{aligned}$$

كما هو متوقع، فإن تقدير β_1 هو بالضبط القيمة المتنبأ بها باستخدام (1.4.18) ولكن لاحظ أن القيمة المقدرة لـ β_0 متحيزة بشكل كبير.

بوجه عام، مقدار التحيز في $\hat{\beta}_1$ يعتمد على β_1، σ^2 و تباين (Y) و u [11]. كما لاحظ Kenneth white el al " هذه هي الفكرة الرئيسية وراء تحيز المعادلات الآنية. على

(11) انظر المعادلة (5.3.18).

العكس حالة نموذج المعادلة الفردية، لايمكننا افتراض أن المتغيرات الموجودة في الجانب الأيمن من المعادلة غير مرتبطة مع مقدار الخطأ"[12].

وتذكر أن هذا التحيز يظل موجوداً حتى في العينات كبيرة الحجم. لاستعراض النتائج الخطيرة المحتملة عند تطبيق OLS في نماذج المعادلات الآنية. هل يوجد اختبار للآنية والذي يمكن الاعتماد عليه لمعرفة ما إذا كانت لدينا المشكلة الخاصة بالمعادلات الآنية أم لا وفقًا للحالة الموجودة لدينا؟ أحد أشكال اختبار Hausman للتخصيص ممكن أن يستخدم لهذا الغرض. وستتم مناقشة ذلك في الفصل (19).

5.18 التلخيص والنتائج : SUMMARY AND CONCLUSIONS

1 - على نقيض نماذج المعادلات الفردية، فإنه في نماذج المعادلات الآنية، يوجد أكثر من متغير تابع واحد أو داخلي، ويوجد عدد لازم من المعادلات مساويًا لعدد المتغيرات الداخلية.

2 - صفة خاصة بنماذج المعادلات الآنية، وهي أن المتغير الداخلي (المنحدر عليه) في المعادلة الواحدة قد يظهر كمتغير مفسر (منحدر) في معادلة أخرى داخل النظام.

3 - كنتيجة لذلك، فإن مثل هذا المتغير المفسر الداخلي العشوائي يكون عادة مرتبطًا مع مقدار الخطأ الموجود في المعادلة التي يظهر فيها كمتغير مفسر.

4 - في مثل هذا الموقف، فإن الـ OLS التقليدية لايمكن تطبيقها، حيث إن المقدرات التي تم الحصول عليها غير متسقة، أي أنها لاتؤول إلى قيم المجتمع الحقيقية حتى إذا كان حجم العينة كبيراً.

5 - مثال Monte Carlo المقدم في الفقرة السابقة، يوضح طبيعة التحيز المرتبط بتطبيق الـ OLS لتقدير المعاملات الخاصة بمعادلة الانحدار عندما يكون المتغير المنحدر مرتبطًا مع مقدار الخطأ، وتلك هي تمامًا الحالة الموجودة في نماذج المعادلات الآنية.

6 - وبما أن نماذج المعادلات الآنية تستخدم باستمرار خصوصًا في النماذج الاقتصادية، هناك بدائل أخرى للتقدير قام العديد من الباحثين باستخدامها. هذه البدائل سيتم استعراضها في الفصل (20) بعد التعرض لموضوع مشكلة التوصيف التي سيتم تناولها في الفصل (13)، وهي منطقيًا موضوع لابد من العرض له قبل الدخول في مشكلة التقدير.

(12) Op.cit., pp.133-134.

EXERCISES تمـــاريـــن :

أسئلة : Questions

1.18 كون نموذج معادلات آنية للعرض والطلب على أطباء الأسنان في الولايات المتحدة الأمريكية. حدد المتغيرات الداخلية والخارجية في النموذج.

2.18 كون نموذجًا بسيطًا للعرض والطلب على المال في الولايات المتحدة، وقارن نموذجك مع النموذج الذي قام به (†)R.Tiegen و (*)K.Brunner, A.H.Meltzer.

3.18 (a) لنموذج العرض والطلب الموجود في مثال 1.18، أوجد صياغة النهاية الاحتمالية لـ $\hat{\alpha}_1$.

(b) ما هي الشروط اللازمة حتى تساوي النهاية الاحتمالية للقيمة الحقيقية α_1؟

4.18 بالنسبة لنموذج IS-LM الذي تم التعرض له سابقًا، أوجد مستوى معدل الفائدة والدخل الذي يتلاءم آنيًا مع توازن المال والسلع في السوق.

5.18 لدراسة العلاقة بين التضخم وأثر ذلك على رأس المال المشترك (‡)Bruno Oudet استخدم النموذج التالي:

$$R_{bt} = \alpha_1 + \alpha_2 R_{st} + \alpha_3 R_{bt-1} + \alpha_4 L_t + \alpha_5 Y_t + \alpha_6 \text{NIS}_t + \alpha_7 I_t + u_{1t}$$

$$R_{st} = \beta_1 + \beta_2 R_{bt} + \beta_3 R_{bt-1} + \beta_4 L_t + \beta_5 Y_t + \beta_6 \text{NIS}_t + \beta_7 E_t + u_{2t}$$

بحيث إن:

L = أساس مالي حقيقي لكل فرد

Y = دخل حقيقي لكل فرد

I = معدل التضخم المتوقع

NIS = متغير مكون حديثًا

E = الفترة المتوقعة لإنهاء السهم، منبية عن الفترة الزمنية المتأخرة لنسب أسعار الأسهم

R_{bt} = دخل السند

R_{st} = عائد الأسهم المشتركة

(*) Some Further Evidence on Supply and Demand Functions for Money," Journal of Finance, vol. 19, May 1964. pp. 240-283.

(†) "Demand and Supply Functions for Money in the United States, "Econometrica, vol. 32, no. 4, October 1964, pp. 476–509.

(‡) Bruno A. Oudet, "The Variation of the Return on Stocks in Periods of Inflation, "Journal of Financial an Quantitative Analysis, vol. 8, no. 2, March 1973, pp. 247–258.

(a) قدم تعليلاً نظريًا لهذا النموذج، وقارن بين أسبابك والأسباب التي قدمها oudet .

(b) ما هي المتغيرات الداخلية في النموذج؟ وما هي المتغيرات الخارجية؟

(c) كيف تعامل قيم R_{bt} في فترات زمنية متأخرة- داخلية أم خارجية؟

6.18 في مقال بعنوان " A model of the distribution of Branded Personal Products in Jamica"[*] . قام Harold J.Lentt و John V.Farley وصمما النموذج التالي (السلع الشخصية تشتمل على كريم الحلاقة، كريم البشرة، مناديل صحية ومعاجين أسنان):

$$Y_{1i} = \alpha_1 + \beta_1 Y_{2i} + \beta_2 Y_{3i} + \beta_3 Y_{4i} + u_{1i}$$

$$Y_{2i} = \alpha_2 + \beta_4 Y_{1i} + \beta_5 Y_{5i} + \gamma_1 X_{1i} + \gamma_2 X_{2i} + u_{2i}$$

$$Y_{3i} = \alpha_3 + \beta_6 Y_{2i} + \gamma_3 X_{3i} + u_{3i}$$

$$Y_{4i} = \alpha_4 + \beta_7 Y_{2i} + \gamma_4 X_{4i} + u_{4i}$$

$$Y_{5i} = \alpha_5 + \beta_8 Y_{2i} + \beta_9 Y_{3i} + \beta_{10} Y_{4i} + u_{5i}$$

بحيث إن:

Y_1 = نسبة المحال التي تخزن المنتج

Y_2 = المبيعات بالوحدة في الشهر

Y_3 = مؤشر الاتصال المباشر بين المستورد والمصنع للمنتج

Y_4 = مؤشر عمليات البيع بالجملة في المنطقة

Y_5 = مؤشر عمق مخزون المنتج (بمعنى متوسط عدد المخزون من المنتج والمتاح في المحال التي تعمل في هذا المنتج)

X_1 = المجتمع المستهدف للمنتج

X_2 = الدخل الفردي في الدائرة التي توجد فيها المنطقة

X_3 = المسافة من مركز المجتمع لـ grairty إلى Kingston

X_4 = المسافة من مركز المجتمع إلى أقرب مدينة للبيع بالجملة

(a) هل يمكنك تحديد المتغيرات الداخلية والخارجية في النموذج السابق؟

(b) هل من الممكن تقدير معادلة أو أكثر في النموذج بطريقة المربعات الصغرى؟ علل إجابتك.

(*) Journal of marketing research, November 1968, pp. 362- 368.

7.18 لدراسة العلاقة بين مصاريف الإعلان ومبيعات السجائر Frank Bass استخدم النموذج التالي(†):

$$Y_{1t} = \alpha_1 + \beta_1 Y_{3t} + \beta_2 Y_{4t} + \gamma_1 X_{1t} + \gamma_2 X_{2t} + u_{1t}$$

$$Y_{2t} = \alpha_2 + \beta_3 Y_{3t} + \beta_4 Y_{4t} + \gamma_3 X_{1t} + \gamma_4 X_{2t} + u_{2t}$$

$$Y_{3t} = \alpha_3 + \beta_5 Y_{1t} + \beta_6 Y_{2t} + u_{3t}$$

$$Y_{4t} = \alpha_4 + \beta_7 Y_{1t} + \beta_8 Y_{2t} + u_{4t}$$

بحيث إن:

Y_1 = لوغاريتم مبيعات السجائر المرشحة (عدد السجائر) مقسوم على المجتمع فوق 20.

Y_2 = لوغاريتم مبيعات السجائر غير المرشحة (عدد السجائر) مقسوم على المجتمع فوق 20 .

Y_3 = لوغاريتم الدعاية بالدولار للسجائر المرشحة مقسوم على المجتمع فوق 20 وعلى مؤشر سعر الدعاية.

Y_4 = لوغاريتم الدعاية بالدولار للسجائر غير المرشحة مقسوم على المجتمع فوق 20 وعلى مؤشر سعر الدعاية.

X_1 = لوغاريتم الدخل الشخصي المتاح مقسوم على المجتمع فوق 20 ومؤشر سعر المستهلك.

X_2 = لوغاريتم سعر السجائر غير المرشحة لكل عبوة مقسوم على مؤشر سعر المستهلك.

(a) في النموذج السابق، الـ Y's متغيرات داخلية، والـ X's خارجية لماذا افترض الباحث أن X_2 متغير خارجي.

(b) إذا تم التعامل مع X_2 على أنه متغير داخلي. كيف يمكنك تعديل النموذج السابق؟

(†) A Simultaneous Equation Regression Study of Advertising and Sales of Cigaretters," Journal of Marketing Research, vol. 6, August 1969, pp. 291–300.

8.18 G.Menges صمم نموذج الاقتصاد القياسي التالي لاقتصاد ألمانيا الغربية(*):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 I_t + u_{1t}$$

$$I_t = \beta_3 + \beta_4 Y_t + \beta_5 Q_t + u_{2t}$$

$$C_t = \beta_6 + \beta_7 Y_t + \beta_8 C_{t-1} + \beta_9 P_t + u_{3t}$$

$$Q_t = \beta_{10} + \beta_{11} Q_{t-1} + \beta_{12} R_t + u_{4t}$$

بحيث إن:

Y = الدخل القومي

I = بنية رأس المال الصافي

C = الاستهلاك الشخصي

Q = الأرباح

P = تكلفة مؤشر المعيشة

R = الإنتاجية الصناعية

t = الزمن

u = الخطأ العشوائي

(a) ما هي المتغيرات التي ستعتبرها داخلية أو التي ستعتبرها خارجية؟

(b) هل هناك أي معادلة في النظام يمكن تقديرها كمعادلة منفردة باستخدام طريقة المربعات الصغرى.

(c) ما هو السبب وراء استخدام المتغير P في دالة الاستهلاك؟

9.18 P.E.Smith و L.E.Gallaway كونا نموذجًا بسيطًا لاقتصاد الولايات المتحدة كالتالي(**):

$$Y_t = C_t + I_t + G_t$$

$$C_t = \beta_1 + \beta_2 \mathrm{YD}_{t-1} + \beta_3 M_t + u_{1t}$$

$$I_t = \beta_4 + \beta_5 (Y_{t-1} - Y_{t-2}) + \beta_6 Z_{t-1} + u_{2t}$$

$$G_t = \beta_7 + \beta_8 G_{t-1} + u_{3t}$$

(*) G. Menges, "Ein Ökonometriches Modell der Bundesrepublik Deutschland (Vier Strukturgleichungen)," I.F.O. Studien, vol. 5, 1959, pp. 1–22.

(**) "A Quarterly Econometric Model of the United States," Journal of American Statistical Association, vol. 56, 1961, pp. 379–383.

بحيث إن :

Y = الإنتاج الكلي القومي

C = مصاريف الاستهلاك الشخصية

I = الاستثمار الكلي المحلي الخاص

G = مصاريف حكومية مضافًا إليها صافي الاستثمار الأجنبي

YD = الدخل المتاح أو الدخل بعد دفع الضريبة

M = المعروض من المال في بداية الفترة الربع سنوية

Z = دخل الملكية بعد الضرائب

t = الزمن

u_1، u_2 و u_3 = أخطاء عشوائية

كل المتغيرات مقاسة في شكل الفرق الأول .

من البيانات الربع سنوية في الفترة 1948–1957 الباحثون طبقوا طريقة المربعات الصغرى لكل معادلة منفردة وحصلوا على النتائج التالية :

$$\hat{C}_t = 0.09 + 0.43\text{YD}_{t-1} + 0.23M_t \qquad R^2 = 0.23$$

$$\hat{I}_t = 0.08 + 0.43(Y_{t-1} - Y_{t-2}) + 0.48Z_t \qquad R^2 = 0.40$$

$$\hat{G}_t = 0.13 + 0.67G_{t-1} \qquad R^2 = 0.42$$

(a) كيف يمكنك تعليل استخدام طريقة المربعات الصغرى لكل معادلة منفردة في مثل هذه الحالة؟

(b) لماذا قيم R^2 منخفضة نسبيًا؟

مسائل : Problems

10.18 جدول (2.18) يعطي بيانات عن Y (الناتج الكلي المحلي)، I (الاستثمار الكلي الخاص المحلي) و C (مصاريف الاستهلاك الشخصية) للولايات المتحدة للفترة 1970–1999 . كل البيانات محسوبة عند 1996 بلايين الدولارات . افترض أن C مرتبطة خطيًا مع Y كما في نموذج Keynesian البسيط لتحديد الدخل الموجود في مثال 2.18 . احصل على تقديرات الـ OLS لمعالم دالة الاستهلاك . خزن النتائج لاستخدامها مرة أخرى لمقارنة تلك الطريقة مع طرق أخرى سيتم مناقشتها في الفصل (20) .

جدول (2.18) مصاريف الاستهلاك الشخصية ، الاستثمار الكلي الخاص المحلي و GDP الولايات المتحدة ، 1970-1999 (بلايين الدولارات)

Observation	C	I	Y	Observation	C	I	Y
1970	2317.5	436.2	3578.0	1985	3820.9	863.4	5717.1
1971	2405.2	485.8	3697.7	1986	3981.2	857.7	5912.4
1972	2550.5	543.0	3998.4	1987	4113.4	879.3	6113.3
1973	2675.9	606.5	4123.4	1988	4279.5	902.8	6368.4
1974	2653.7	561.7	4099.0	1989	4393.7	936.5	6591.9
1975	2710.9	462.2	4084.4	1990	4474.5	907.3	6707.9
1976	2868.9	555.5	4311.7	1991	4466.6	829.5	6676.4
1977	2992.1	639.4	4511.8	1992	4594.5	899.8	6880.0
1978	3124.7	713.0	4760.6	1993	4748.9	977.9	7062.6
1979	3203.2	735.4	4912.1	1994	4928.1	1107.0	7347.7
1980	3193.0	655.3	4900.9	1995	5075.6	1140.6	7543.8
1981	3236.0	715.6	5021.0	1996	5237.5	1242.7	7813.2
1982	3275.5	615.2	4913.3	1997	5423.9	1393.3	8159.5
1983	3454.3	673.7	5132.3	1998	5678.7	1566.8	8515.7
1984	3640.6	871.5	5505.2	1999	5978.8	1669.7	8875.8

لاحظ أن : C = مصاريف الاستهلاك الشخصي
I = الاستثمار الكلي الخاص المحلي
Y = الناتج الكلي المحلي

المصدر : Economic Report of the president, 2001, table B-2, p.276

11.18 باستخدام البيانات المعطاة في تمرين 10.18 ، قم بانحدار الاستثمار الكلي المحلي I على GDP ، وخزن النتائج لاستخدامات أخرى في الفصل التالي .

12.18 اعتبر تعريف الاقتصاد الكلي التالي :

$$C + I = Y \qquad (=\text{GDP})$$

كما سبق ، افترض التالي :

$$C_t = \beta_0 + \beta_1 Y_t + u_t$$

واعتبر النموذج المسرع للاقتصاد الكلي التالي :

$$I_t = \alpha_0 + \alpha_1 (Y_t - Y_{t-1}) + v_t$$

بحيث إن u ، v مقادير الأخطاء . من البيانات الموجودة في تمرين 10.18 ، قدر النموذج المسرع ، واحتفظ بالنتائج لدراسة أخرى .

الفصل التاسع عشر

مشـــكلة التـــوصيـــف

THE IDENTIFICATION PROBLEM

في هذا الفصل، نتناول طبيعة وأهمية مشكلة التوصيف. أساس مشكلة التوصيف كالتالي: تذكر نموذج العرض والطلب المقدم في فقرة 2.18. افترض أن لدينا بيانات سلاسل زمنية على Q و P فقط بدون أي معلومات إضافية (مثل دخل المستهلك، الأسعار السائدة في الفترة السابقة والظروف المناخية).

مشكلة التوصيف تتمثل في الإجابة عن السؤال التالي: بمعلومية بيانات P و Q فقط، كيف لك أن تعرف إذا كنا نقدر دالة العرض أو دالة الطلب؟ على العكس إذا كنا نوفق دالة طلب، كيف يمكن أن نضمن أن ما نقوم به في الحقيقة هو تقدير دالة الطلب وليس شيئًا أخر؟

بمجرد التفكير اللحظي في إجابة السؤال السابق، سنجد أنه من الضروري معرفة ذلك قبل البدء في تقدير معلمات دالة الطلب. في هذا الفصل، سنوضح كيفية حل مشكلة التوصيف. سنقدم في البداية بعض الرموز والتعريفات، ثم نشرح مشكلة التوصيف بأمثلة عديدة.

ثم نتبع ذلك بالقواعد المستخدمة لمعرفة ما إذا كانت معادلة ما داخل نموذج معادلات آنية موصفة أم لا، بمعنى معرفة ما إذا كانت العلاقة التي نرغب في تقديرها بالفعل هي دالة عرض أو دالة طلب أو شيء آخر.

1.19 رموز وتعريفات : NOTATIONS AND DEFINITIONS

لتسهيل مناقشتنا، دعنا نقدم الرموز والتعريفات التالية: النموذج العام لـ M من المعادلات في M متغير داخلي، أو M متغيرات تابعة تبادلية، يمكن كتابته كما في المعادلة (1.1.19) كالتالي:

$$
\begin{aligned}
Y_{1t} &= \qquad\quad \beta_{12}Y_{2t} + \beta_{13}Y_{3t} + \cdots + \beta_{1M}Y_{Mt} \\
&\qquad + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + \cdots + \gamma_{1K}X_{Kt} + u_{1t} \\
Y_{2t} &= \beta_{21}Y_{1t} \qquad + \beta_{23}Y_{3t} + \cdots + \beta_{2M}Y_{Mt} \\
&\qquad + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + \cdots + \gamma_{2K}X_{Kt} + u_{2t} \\
Y_{3t} &= \beta_{31}Y_{1t} + \beta_{32}Y_{2t} \qquad + \cdots + \beta_{3M}Y_{Mt} \\
&\qquad + \gamma_{31}X_{1t} + \gamma_{32}X_{2t} + \cdots + \gamma_{3K}X_{Kt} + u_{3t} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
Y_{MT} &= \beta_{M1}Y_{1t} + \beta_{M2}Y_{2t} + \cdots + \beta_{M,M-1}Y_{M-1,t} \\
&\qquad + \gamma_{M1}X_{1t} + \gamma_{M2}X_{2t} + \cdots + \gamma_{MK}X_{Kt} + u_{Mt}
\end{aligned}
\qquad (1.1.19)
$$

بحيث إن:

Y_1، Y_2، . . . ، $Y_m = M$ متغيرات داخلية أو تابعة تبادلية.

X_1، X_2، . . . ، $X_K = K$ متغيرات سابقة التحديد (واحد من هذه المتغيرات X، يمكن أن يأخذ قيمة الوحدة للسماح بوجود الجزء المقطوع من المحور الصادي في كل معادلة).

u_1، u_2، . . . ، $u_M = M$ مقادير تشتت (خطأ) عشوائية

$t = 1, 2, ..., T$ = العدد الكلي للمشاهدات

β's = معاملات المتغيرات الداخلية

γ's = معاملات المتغيرات المحددة سابقًا

وكملحوظة عابرة، لاحظ أنه ليس بالضرورة ظهور كل متغير في كل معادلة. وفي واقع الأمر، كما سنرى في الفقرة 2.19 يجب ألا يحدث إذا كانت هناك إمكانية لتوصيف المعادلة.

كما يتضح من المعادلة (1.1.19) المتغيرات الموجودة في نموذج المعادلات الآنية لها نوعان: - داخلية، بمعنى هؤلاء (قيم هذه المتغيرات) محددة من خلال النموذج ومتغيرات سابقة التحديد، بمعنى (قيم المتغيرات) محددة خارج النموذج. المتغيرات الداخلية ينظر لها على أنها عشوائية، في حين المتغيرات المحددة سابقًا تعامل على أنها غير عشوائية.

المتغيرات المحددة سابقًا يتم تقسيمها لفئتين : خارجية، تشمل القيم الحالية والمتأخرة، وداخلية في فترات زمنية متأخرة. بمعنى X_{1t} هو قيمة حالية (في الزمن الحالي) لمتغير خارجي في حين $X_{1(t-1)}$ هي قيمة لمتغير خارجي، ولكن في فترة زمنية متأخرة، وحدة زمنية واحدة متأخرة Y_{t-1} هي قيمة لمتغير داخلي في فترة زمنية متأخرة بوحدة زمنية واحدة، وينظر لها على أنها غير عشوائية، وبالتالي هي متغير محدد سابقًا[1].

باختصار، المتغيرات الخارجية الحالية والخارجية في فترات زمنية متأخرة، والمتغيرات الداخلية في فترات زمنية متأخرة، يتم التعامل معها على أنها متغيرات محددة سابقًا، بمعنى أن قيمها لم تحدد من خلال النموذج في الفترة الزمنية الحالية.

سيرجع إلى الباحث أن يحدد أي المتغيرات داخلية، وأي منها متغيرات محددة سابقًا، على الرغم من أن المتغيرات (غير الاقتصادية) مثل درجة الحرارة وسقوط الأمطار واضح إنها متغيرات خارجية أو محددة سابقًا. فعلى الباحث أن يبذل مجهودًا كبيرًا في التقسيم الجيد للمتغيرات الاقتصادية كداخلية أو محددة سابقًا. كما على الباحث أن يحدد الخلفية السابقة، أو النظرية لأساس تقسيم المتغيرات. عمومًا مؤخرًا في هذا الفصل، سنعطي اختبارًا إحصائيًا لخارجية النشأة.

المعادلة المعطاة في (1.1.19) معرفة على أنها معادلات بنائية أو سلوكية، حيث إنها توصف بناء (لنموذج اقتصادي) اقتصاد ما أو سلوك مؤسسة اقتصادية (بمعنى المستهلك أو المنتج). الـ βs والـ γ's تعرف على أنها معلمات بنائية أو معاملات.

من المعادلات البنائية، يمكن أن يتم الحل للحصول على قيم المتغيرات الداخلية M، والحصول على المعادلات المخفضة، والمعاملات المخفضة المرتبطة بها. المعادلة في الشكل المخفض يتم فيها تمثيل المتغير الداخلي في شكل متغيرات محددة سابقًا فقط والخطأ العشوائي. للتوضيح اعتبر نموذج Keynesian لتحديد الدخل المعطى في الفصل (18):

(3.2.18) دالة الاستهلاك : $0 < \beta_1 < 1$ $\quad C_t = \beta_0 + \beta_t Y_t + u_t$

(4.2.18) وحدة الدخل : $Y_t = C_t + I_t$

(1) من المفترض بشكل ضمني هنا أن مقدار الخطأ العشوائي، u's، غير مرتبط تسلسليًا، وإذا لم تكن تلك هي الحالة الموجودة فإن Y_{t-1} سيكون مرتبطًا مع القيمة الحالية لمقدار الخطأ u_t وبالتالي لايمكن التعامل معه على أنه متغير محدد سابقًا.

في هذا النموذج C (الاستهلاك)، وY (الدخل) متغيرات داخلية، وI (مصاريف الاستثمار)، يتم التعامل معها على أنها متغير خارجي.

كل من هاتين المعادلتين تمثلان معادلات بنائية، (4.2.18) تعتبر معادلة الوحدة. كالعادة الـ MPC β_1 مفترض أن يقع بين 0 و 1. إذا عوضنا عن (3.2.18) في (4.2.18) ويعمل بعض الخطوات الجبرية البسيطة نحصل على:

$$Y_t = \Pi_0 + \Pi_1 I_t + w_t \qquad \textbf{(2.1.19)}$$

حيث:

$$\Pi_0 = \frac{\beta_0}{1-\beta_1}$$

$$\Pi_1 = \frac{1}{1-\beta_1} \qquad \textbf{(3.1.19)}$$

$$w_t = \frac{u_t}{1-\beta_1}$$

معادلة (2.1.19) تعتبر معادلة في الشكل المخفض، حيث تعبر عن المتغير الداخل Y، وحدة كدالة في متغير خارجي (سابق التحديد)، I ومقدار الخطأ العشوائي u، Π_0 و Π_1 معاملات مخفضة مرتبطة بمتغيرات النموذج. لاحظ أن هذه المعاملات المخفضة هي عبارة عن توليفة غير خطية من المعامل أو المعاملات البنائية. بالتعويض عن قيمة Y من (2.1.19) إلى C في (3.2.18) نحصل على معادلة مخفضة أخرى:

$$C_t = \Pi_2 + \Pi_3 I_t + w_t \qquad \textbf{(4.1.19)}$$

بحيث إن:

$$\Pi_2 = \frac{\beta_0}{1-\beta_1} \qquad \Pi_3 = \frac{\beta_1}{1-\beta_1}$$

$$w_t = \frac{u_t}{1-\beta_1}$$

المعاملات المخفضة مثل Π_1 و Π_3 معروفة أيضًا باسم التأثير أو المضاريب قصيرة الأجل، حيث إنها تقيس التأثير الفوري على المتغير الداخلي للتغيير بوحدة واحدة في قيمة المتغير الخارجي[2].

(2) في نماذج الاقتصاد القياسي، المتغيرات الخارجية تلعب دوراً مهمًا. فغالبًا مثل هذه المتغيرات تكون تحت تصرف الحكومة. أمثلة على ذلك معدل الضرائب الشخصية والمشتركة، العوض عن البطالة، subsidies وإلى ما غير ذلك.

في نموذج Keynesian السابق إذا زادت مصاريف الاستثمار بواحد دولار مثلاً، وإذا فرضنا أن MPC يساوي 0.8 إذن من (3.1.19) نحصل على $\Pi_1 = 5$. هذه النتيجة تعني أن زيادة الاستثمار بدولار واحد سيؤدي مباشرة (أي في الفترة الحالية) إلى زيادة في الدخل بمقدار 5 دولارات، أي زيادة بخمس مرات. بالمثل، تحت الشروط المفروضة، فإن (5.1.19) توضح أن $\Pi_3 = 4$ يعني أن لكل زيادة 1 دولار في مصاريف الاستثمار ستؤدي مباشرة إلى زيادة 4 دولارات في مصاريف الاستهلاك.

في **مجال الاقتصاد القياسي**، معادلات مثل (4.2.18) أو $Q_t^d = Q_t^s$ (الكمية المطلوبة تساوي الكمية المعروضة) معروفة باسم شرط التوازن. معادلة (4.2.18) تعني أن الدخل التجميعي Y لابد أن يساوي الاستهلاك التجميعي (بمعنى مصاريف الاستهلاك بالإضافة إلى مصاريف الاستثمار) عندما يحدث التوازن، المتغيرات الداخلية تفترض قيمها التوازنية[3]. لاحظ صفة شيقة في المعادلات المخفضة. بما أن المتغيرات المحددة سابقًا، والأخطاء العشوائية تظهر فقط في الجانب الأيمن من هذه المعادلات، وبما أن المتغيرات المحددة سابقًا مفترض أن تكون غير مرتبطة مع مقادير الأخطاء، فإن طريقة OLS من الممكن تطبيقها لتقدير معاملات المعادلات المخفضة (الـ Π's). من المعاملات المخفضة المقدرة، يمكن للباحث أن يقدر المعاملات البنائية (الـ βs) كما سيوضح لاحقًا. هذه الطريقة معروفة باسم طريقة المربعات الصغرى غير المباشرة (ILS)، والمعاملات البنائية المقدرة تسمى مقدرات OLS.

سندرس طريقة ILS بمزيد من التفاصيل في الفصل (20). في الوقت الحالي، لاحظ أنه بما أن المعاملات المخفضة من الممكن تقديرها باستخدام طريقة OLS. وبما أن هذه المعاملات توليفات من المعاملات البنائية، فهناك احتمال أن المعاملات البنائية يمكن استرجاعها من المعاملات المخفضة، وتقدير المعاملات البنائية من الأشياء المهمة التي يحتاج إليها الباحث. كيف يمكن للفرد أن يسترجع المعاملات البنائية من المعاملات المخفضة؟ الإجابة معطاة في فقرة 2.19، الإجابة التي ستطرح نقاشًا مهمًا حول أساس مشكلة التوصيف.

---

(3) للمزيد من التفاصيل، اقرأ

Jan Kmenta, Elements of Econometrics, 2d ed., Macmillan, New York, 1986, pp. 723–731.

2.19 مشكلة التوصيف : THE IDENTIFICATION PROBLEM

مشكلة التوصيف تعني ما إذا كان من الممكن الحصول على تقدير رقمي لمعاملات المعادلة البنائية من المعاملات المقدرة المخفضة أم لا. إذا كان من الممكن عمل ذلك، فإننا نقول إن هذه المعادلة يمكن توصيفها. إذا لم يكن من الممكن عمل ذلك، فبالتالي نقول إن المعادلة محل الاهتمام لا يمكن توصيفها أو موصفة بأقل مما يجب.

المعادلة الموصفة، إما أن تكون تامة التوصيف (أو كاملة التوصيف أو موصفة فقط) أو موصفة بأكثر مما يجب. فيقال إنها معادلة تامة التوصيف، إذا كانت هناك قيم رقمية وحيدة للمعلمات البنائية يمكن الحصول عليها. ويقال معادلة موصفة بأكثر مما يجب إذا كان هناك أكثر من قيمة رقمية يمكن الحصول عليها لبعض معلمات المعادلات البنائية. الظروف الخاصة بكل حالة من هذه الحالات سيتم استعراضها في الفقرة التالية.

مشكلة التوصيف تظهر بسبب وجود أكثر من مجموعات من المعاملات البنائية التي يمكن استخدامها لنفس مجموعة البيانات. بعبارة أخرى، المعادلة المخفضة المعطاة من الممكن أن تتماشى مع معادلات بنائية مختلفة أو فروض (نماذج) مختلفة، ومن الصعب تحديد أي فرض (نموذج) تتم دراسته. في المتبقي من هذه الفقرة، سنستعرض العديد من الأمثلة لتوضيح طبيعة مشكلة التوصيف.

التوصيف بأقل مما يجب : Underidentification

اعتبر مرة أخرى نموذجي العرض والطلب (1.2.18) و(2.2.18)، مع شروط توازن أو توافق السوق الخاصة، فإن الطلب يساوي العرض، فمن خلال شرط التوازن نحصل على:

$$\alpha_0 + \alpha_1 P_t + u_{1t} + \beta_0 + \beta_1 P_t + u_{2t} \qquad (1.2.19)$$

بحل (1.2.19) نحصل على السعر التوازني التالي:

$$P_t = \Pi_0 + v_t \qquad (2.2.19)$$

بحيث إن:

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} \qquad (3.2.19)$$

$$v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1} \qquad (4.2.19)$$

وبالتعويض عن P_t من (2.2.19) في (1.2.18) أو (2.2.18) نحصل على الكمية التوازنية التالية:

$$Q_t = \Pi_1 + w_t \qquad (5.2.19)$$

بحيث إن:

$$\Pi_1 = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1} \qquad (6.2.19)$$

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1} \qquad (7.2.19)$$

لاحظ أن الأخطاء v_t و w_t ما هي إلا توليفة خطية من مقادير الأخطاء الأصلية u_1 و u_2.

المعادلات (2.2.19) و (5.2.19) هي معادلات مخفضة الشكل. الآن نموذجنا الخاص بالعرض والطلب يشتمل على أربع معاملات بنائية α_0، α_1، β_0 و β_1 ولكن لاتوجد طريقة وحيدة لتقديرها. لماذا؟ الإجابة موجودة في المعاملين المخفضين الموجودين في (3.2.19) و(6.2.19). هذه المعاملات المخفضة تشمل كل المعلمات الأربع البنائية ولكن لاتوجد طريقة لتقدير الأربعة مجاهيل البنائية من المعاملين المخفضين الاثنين فقط. تذكر في المرحلة الثانوية في مادة الجبر، عرفنا أن تقدير أربعة مجاهيل لابد أن يكون من خلال أربع معادلات (مستقلة)، وعمومًا لتقدير k مجهول لابد أن يكون لدينا k معادلة (مستقلة).

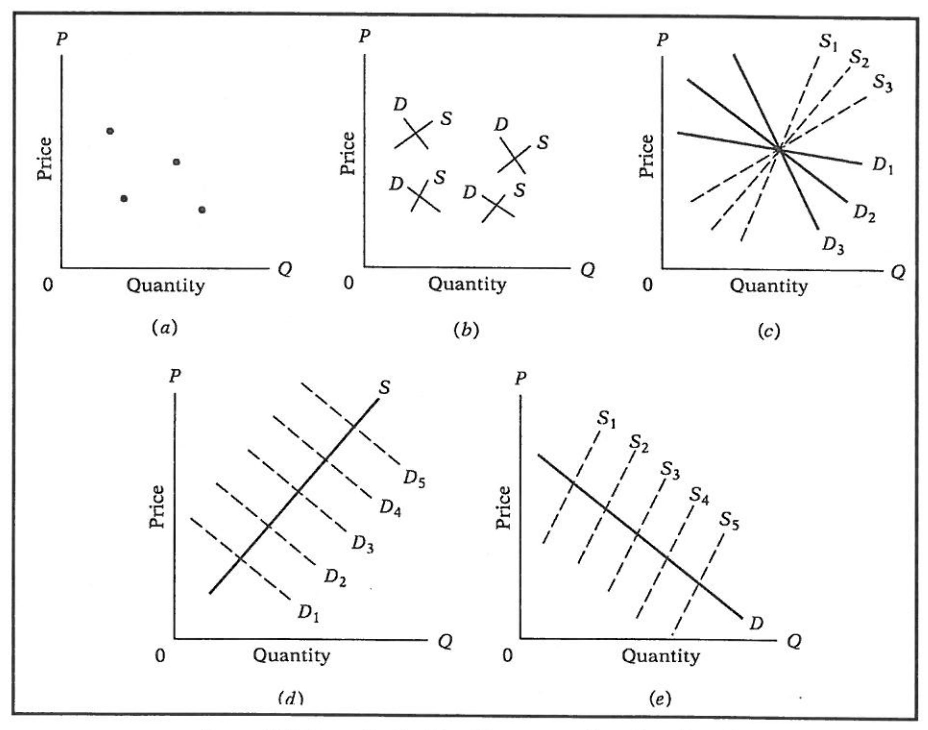
وبشكل عارض، إذا قمنا بالانحدار المخفض (2.2.19) و (5.2.19) سنرى عدم وجود متغيرات مفسرة، فالموجود هو الثابت فقط، وهذه الثوابت ستعطي ببساطة القيمة المتوسطة لـ P و Q (لماذا؟).

كل ما يعنيه ذلك، هو أنه بمعلومية بيانات سلسلة زمنية عن P (السعر) و Q (الكمية) وعدم وجود أي معلومات أخرى إضافية، لاتوجد طريقة يستطيع الباحث من خلالها ضمان ما إذا كان يقدر دالة طلب أم دالة عرض. بمعنى أن القيم المعطاة لكل من P_t و Q_t تمثل نقطة تقاطع بين منحنى العرض ومنحنى الطلب، حيث إن شرط التوازن يتطلب أن يتساوى العرض مع الطلب. لرؤية ذلك بوضوح انظر شكل الانتشار المعطى في شكل (1.19).

شكل (a1.19) يعطي القليل من نقط الانتشار التي تربط Q مع P. كل نقطة انتشار تمثل تقاطعًا بين منحنى الطلب ومنحنى العرض، كما هو موضح في شكل

($b1.19$) لاتوجد طريقة أكيدة لمعرفة أي منحنى عرض وطلب من هذه العائلة من المنحنيات تنتمي إليه هذه النقطة. بالطبع نحن نحتاج إلى المزيد من المعلومات عن طبيعة منحنيات العرض والطلب.

على سبيل المثال، إذا تحرك منحنى الطلب مع الزمن بسبب تغير في الدخل، الذوق وإلى ما غير ذلك، ولكن ظل منحنى العرض كما هو كما في شكل ($d1.19$) فإن نقاط الإنتشار تحدد منحنى عرض، وفي مثل هذه الحالة، فإننا نقول إن منحنى العرض منحنى موصف (معرف). وبنفس الطريقة إذا تحرك منحنى العرض مع الزمن بسبب تغييرات في الظروف المناخية (في حالة السلع الزراعية) أو بسبب أي عوامل أخرى، ولكن ظل منحنى الطلب كما هو فإنه كما في شكل ($e1.19$) فإن نقاط الانتشار تحدد منحنى الطلب. في مثل هذه الحالة نقول إن منحنى الطلب منحنى موصف.



شكل (1.19) دوال عرض وطلب افتراضية ومشكلة التوصيف

هناك طريقة بديلة، وربما أكثر وضوحًا لفهم مشكلة التوصيف. افترض أننا ضربنا المعادلة (1.2.18) في λ ($0 \le \lambda \le 1$) والمعادلة (2.2.18) في $1-\lambda$ للحصول على المعادلات التالية (لاحظ أننا أسقطنا الترميز على Q):

$$\lambda Q_t = \lambda\alpha_0 + \lambda\alpha_1 P_t + \lambda u_{1t} \quad (8.2.19)$$

$$(1-\lambda)Q_t = (1-\lambda)\beta_0 + (1-\lambda)\beta_1 P_t + (1-\lambda)u_{2t} \quad (9.2.19)$$

بإضافة هاتين المعادلتين نحصل على التوليفة الخطية التالية للمعادلات الأصلية للعرض والطلب. :

$$Q_t = \gamma_0 + \gamma_1 P_t + w_t \quad (10.2.19)$$

بحيث إن :

$$\gamma_0 = \lambda\alpha_0 + (1-\lambda)\beta_0$$
$$\gamma_1 = \lambda\alpha_1 + (1-\lambda)\beta_1 \quad (11.2.19)$$
$$w_t = \lambda u_{1t} + (1-\lambda)u_{2t}$$

المعادلة الهجين أو المختلفة (10.2.19) لايمكن بمجرد الملاحظة تفريقها عن (1.2.18) أو (2.2.18) حيث إنها تشتمل على انحدار Q و P. بالتالي إذا كان لدينا بيانات سلاسل زمنية عن P و Q فقط فإن أياً من (1.2.18)، (2.2.18) أو (10.2.19) يمكن استخدامها على نفس البيانات. بمعنى آخر، نفس البيانات يمكن التعامل معها بالفروض (1.2.18)، (2.2.18) أو (10.2.19) ولاتوجد طريقة لتحديد أي من هذه الفروض يخضع للدراسة والاختبار.

حتى تكون المعادلة موصفة ، أي أن معلمات المعادلة يمكن تقديرها، لابد من توضيح أنه أي مجموعة ما معطاة من البيانات لن تكون معادلة بنائية مماثلة في الشكل للمعادلة التي نرغب فيها. فإذا أردنا تقدير دالة الطلب ، يجب أن نوضح أن البيانات المعطاة غير متسقة مع دالة العرض، أو مع أي معادلة أخرى هجين أو خليط بين العرض والطلب.

تامة التوصيف أو موصفة فقط : Just, or Exact, Identification

السبب الذي جعلنا لانستطيع توصيف دالة الطلب السابقة، أو دالة العرض، هو ظهور نفس المتغيرات P و Q في كل من المعادلتين، ولاتوجد أي معلومات إضافية كالمعطاة في شكل (d1.19) أو e. لكن دعنا نعتبر نموذج العرض والطلب التالي :

(12.2.19) دالة الطلب : $\alpha_1 < 0, \alpha_2 > 0$ $\quad Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$

(13.2.19) دالة العرض : $\beta_1 > 0$ $\quad Q_t = \beta_0 + \beta_1 P_t + u_{2t}$

بحيث إن I = دخل المستهلك، متغير خارجي، وكل المتغيرات معرفة كما سبق تعريفها من قبل.

لاحظ أن الفرق الوحيد بين النموذج السابق ونموذجنا الأصلي الخاص بالعرض والطلب، هو موجود متغير إضافي في دالة الطلب وهو متغير الدخل. من النظرية الاقتصادية للطلب، نعرف أن الدخل عادة يعتبر محدداً مهمًا للطلب على معظم السلع والخدمات. وبالتالي وجوده في دالة الطلب سيعطينا المزيد من المعلومات عن سلوك المستهلك. للعديد من السلع من المتوقع أن يكون للدخل تأثير إيجابي (موجب) على الاستهلاك ($\alpha_2 > 0$).

باستخدام طريقة توازن السوق، الكمية المطلوبة= الكمية المعروضة، نحصل على:

$$\alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t} = \beta_0 + \beta_1 P_t + u_{2t} \qquad (14.2.19)$$

وبحل المعادلة (14.2.19) نحصل على القيمة التوازنية التالية لـ P_t:

$$P_t = \Pi_0 + \Pi_1 I_t + v_t \qquad (15.2.19)$$

حيث إن المعاملات المخفضة الشكل هي:

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}$$

$$\Pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1} \qquad (16.2.19)$$

و

$$v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1}$$

بالتعويض عن القيمة التوازنية لـ P_t في معادلة الطلب أو العرض السابقة، نحصل على الكمية التوازنية التالية:

$$Q_t = \Pi_2 + \Pi_3 I_t + w_t \qquad (17.2.19)$$

بحيث إن:

$$\Pi_2 = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1}$$

$$\Pi_3 = -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1} \qquad (18.2.19)$$

و

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1}$$

وبما أن (15.2.19) و (17.2.19) كل منهما معادلة مخفضة الشكل، فإن طريقة OLS يمكن تطبيقها لتقدير معلماتهم. الآن نموذج العرض والطلب (12.2.19) و(13.2.19) يحتوى خمس معاملات بنائية α_0 و α_1 و α_2 و β_1 و β_2 ولكن توجد أربع معادلات فقط لتقديرها وهي المعاملات الأربعة المخفضة Π_0، Π_1، Π_2، Π_3 المعطاة في (16.2.19) و (18.2.19).

وبالتالي، فلايمكن الحصول على حل وحيد للمعاملات البنائية، ولكن من السهل توضيح أن معاملات دالة العرض من الممكن توصيفها (تقديرها) حيث إن:

$$\begin{aligned} \beta_0 &= \Pi_2 - \beta_1 \Pi_0 \\ \beta_1 &= \frac{\Pi_3}{\Pi_1} \end{aligned} \qquad (19.2.19)$$

ولكن لاتوجد طريقة وحيدة لتقدير معلمات دالة الطلب، وبالتالي تظل موصفة بأقل مما يجب. وبشكل عابر، لاحظ أن المعامل البنائي β_1 هو دالة غير خطية في المعاملات المخفضة الشئ الذي ينتج عنه بعض المشاكل الخاصة بتقدير الأخطاء القياسية للمقدر β_1، كما سنرى في الفصل (20).

لإثبات أن دالة الطلب (12.2.19) لايمكن توصيفها (تقديرها). دعنا نضرب هذه المعادلة في $\lambda (0 \leq \lambda \leq 1)$ و (13.2.19) نضربها في $1-\lambda$ ونجمع المعادلتين معًا، فنحصل على المعادلة المختلفة التالية:

$$Q_t = \gamma_0 + \gamma_1 P_t + \gamma_2 I_t + w_t \qquad (20.2.19)$$

بحيث إن:

$$\begin{aligned} \gamma_0 &= \lambda\alpha_0 + (1-\lambda)\beta_0 \\ \gamma_1 &= \lambda\alpha_1 + (1-\lambda)\beta_1 \\ \gamma_2 &= \lambda\alpha_2 \end{aligned} \qquad (21.2.19)$$

و

$$w_t = \lambda u_{1t} + (1-\lambda)u_{2t}$$

معادلة (20.2.19) لايمكن تفريقها بمجرد النظر عن دالة الطلب (12.2.19) ولكن يمكن تفريقها عن دالة العرض (13.2.19) والتي لاتحتوي على المتغير I كمتغير مفسر، وبالتالي، فإن دالة الطلب تظل غير موصفة.

لاحظ حقيقة مثيرة: وجود متغير إضافي في دالة الطلب جعلنا قادرين على توصيف دالة العرض ! لماذا؟ اشتمال دالة الطلب على متغير الدخل، يعطينا بعض المعلومات الإضافية عن تنوع الدالة، كما هو موضح في شكل ($d1.19$) الشكل يوضح كيفية تفاعل (تقاطع) منحنى العرض الثابت مع منحنى الطلب المتحرك (بالنسبة للتغير في الدخل) مما يجعلنا قادرين على رسم (توصيف) منحنى العرض. وكما سنرى بعد قليل، في كثير من الأحيان يعتمد توصيف المعادلة على ما إذا كانت تحتوي على واحد أو أكثر من المتغيرات الأخرى والموجودة في معادلات أخرى في النموذج.

ولكن دعنا نفترض النموذج التالي للعرض والطلب :

(12.2.19) دالة الطلب: $\alpha_1 < 0, \alpha_2 > 0$ $\quad Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$

(22.2.19) دالة العرض: $\beta_1 > 0, \beta_2 > 0$ $\quad Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$

حيث إن دالة الطلب تظل كما سبق، ولكن دالة العرض تحتوي على متغير مفسر إضافي هو السعر في فترة زمنية واحدة سابقة (متأخرة). دالة العرض تفترض أن الكمية المعروضة من سلعة ما تعتمد على السعر الحالي والسعر في الفترة السابقة. هذا النموذج عادة يستخدم لشرح عرض الكثير من السلع الزراعية. لاحظ أن P_{t-1} هي متغير محدد سابقًا، حيث إن قيمته معروفة عند الوقت t.

باستخدام طريقة توازن السوق نجد أن :

$$\alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t} = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t} \qquad (23.2.19)$$

وبحل هذه المعادلة نحصل على السعر التوازني التالي :

$$P_t = \Pi_0 + \Pi_1 I_t + \Pi_2 P_{t-1} + v_t \qquad (24.2.19)$$

بحيث إن :

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}$$

$$\Pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1} \qquad (25.2.19)$$

$$\Pi_2 = \frac{\beta_2}{\alpha_1 - \beta_1}$$

و

$$v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1}$$

بالتعويض عن السعر التوازني في معادلة العرض والطلب، نحصل على الكمية المناظرة التوازنية التالية:

$$Q_t = \Pi_3 + \Pi_4 I_t + \Pi_5 P_{t-1} + w_t \qquad (26.2.19)$$

حيث إن المعاملات المخفضة هي :

$$\Pi_3 = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1}$$

$$\Pi_4 = -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1} \qquad (27.2.19)$$

$$\Pi_5 = \frac{\alpha_1\beta_2}{\alpha_1 - \beta_1}$$

و

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1}$$

نموذج العرض والطلب المعطى في المعادلات (12.2.19) و (22.2.19) يحتوي على ست معاملات بنائية- α_1 و α_2 و α_3 و β_0 و β_1 و β_2 وهناك ست معاملات مخفضة- Π_0، Π_1، Π_2، Π_3، Π_4 و Π_5 لتقديرها. وبالتالي لدينا ست معادلات وستة مجاهيل، وبالتالي من الممكن وجود مقدرات وحيدة. وبالتالي معاملات كل من معادلات العرض والطلب يمكن توصيفها، والنظام ككل يمكن توصيفه (في تمرين 2.19 يسأل القارئ أن يعبر عن الست معاملات البنائية في صورة الست معاملات المخفضة السابق ذكرها لإثبات أنه يمكن الحصول على تقدير وحيد للنموذج).

للتأكد من أن دوال العرض والطلب السابقة يمكن توصيفها. يمكن أن نلجأ إلى فكرة ضرب دالة الطلب (12.2.19) في λ ($0 \le \lambda \le 1$) ودالة العرض (22.2.19) في ($1-\lambda$) ونجمعهما للحصول على المعادلة المختلفة. هذه المعادلة المختلفة ستحتوي على كل من المتغيرين السابق تحديدهما I_t و P_{t-1} وبالطبع ستختلف عن دالة العرض، وعن دالة الطلب أيضًا حيث إن الأولى لاتحتوي على P_{t-1} والأخيرة لاتحتوي على I_t.

التوصيف بأكثر مما يجب : Overidentification

دخل أو ثروة المستهلك تعتبر محدداً مهمًا للطلب على بعض السلع والخدمات، وبالتالي دعنا نعرف دالة الطلب (12.2.19) كالتالي، وستظل دالة العرض كما هي معرفة من قبل:

(28.2.19) دالة الطلب:
$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t}$$

(22.2.19) دالة العرض:
$$Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$$

حيث تم إضافة المتغير R الذي يعبر عن الثروة، لمعظم السلع والخدمات، فإن الثروة مثل الدخل من المتوقع أن يكون لها تأثير إيجابي (موجب) على الاستهلاك.

لمساواة العرض مع (الطلب)، نحصل على السعر والكمية التوازنية التالية:

$$P_t = \Pi_0 + \Pi_1 I_t + \Pi_2 R_t + \Pi_3 P_{t-1} + v_t \qquad (29.2.19)$$

$$Q_t = \Pi_4 + \Pi_5 I_t + \Pi_6 R_t + \Pi_7 P_{t-1} + w_t \qquad (30.2.19)$$

بحيث إن:

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} \qquad \Pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1}$$

$$\Pi_2 = -\frac{\alpha_3}{\alpha_1 - \beta_1} \qquad \Pi_3 = \frac{\beta_2}{\alpha_1 - \beta_1}$$

$$\Pi_4 = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1} \qquad \Pi_5 = -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1} \qquad (31.2.19)$$

$$\Pi_6 = -\frac{\alpha_3\beta_1}{\alpha_1 - \beta_1} \qquad \Pi_7 = \frac{\alpha_1\beta_2}{\alpha_1 - \beta_1}$$

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1} \qquad v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1}$$

نموذج العرض والطلب السابق يحتوي على سبع معاملات بنائية، ولكن هناك ثماني معادلات لتقديرها- المعاملات المخفضة الثماني المعطاة في (31.2.19). وبالتالي فإن عدد المعادلات أكبر من عدد المجاهيل، وكنتيجة لذلك، لا يمكن الحصول على تقدير وحيد للمعلمات، ويمكن إثبات ذلك بسهولة. من المعاملات السابقة المخفضة نحصل على:

$$\beta_1 = \frac{\Pi_6}{\Pi_2} \qquad (32.2.19)$$

أو

$$\beta_1 = \frac{\Pi_5}{\Pi_1} \qquad (33.2.19)$$

وبالتالي، هناك تقديرات لمعامل السعر في دالة العرض، ولايوجد ما يضمن

تطابق هاتين القيمتين أو الحلين معاً[4]. الأكثر من ذلك، بما أن β_1 تظهر في مقام كل المعاملات المخفضة، فإن الغموض أو الالتباس في تقدير β_1 سينتقل إلى باقي المقدرات أيضًا.

لماذا تم توصيف دالة العرض في النظام (12.2.19) و (22.2.19) ولكن ليس في النظام (18.2.19) و (22.2.19). على الرغم من أنه في كلتا الحالتين، دالة العرض ظلت كما هي؟ الإجابة أن لدينا "الكثير" أو أكثر من اللازم من المعلومات لتوصيف منحنى العرض. هذه الحالة هي الحالة العكسية للتوصيف بأقل مما يجب، حيث يكون لدينا معلومات أقل من المطلوب. وجود معلومات أكثر من اللازم ناتج من حقيقة إبعاد متغير الدخل عن دالة العرض، كما في نموذج (12.2.19) و (22.2.19) مما كان كافيًا لتوصيف المعادلة ولكن في النموذج (28.2.19) و (22.2.19) فإن دالة العرض لاتستبعد فقط متغير الدخل، ولكن أيضًا متغير الثروة، وبمعنى آخر في الحالة الأخيرة نحن نضع قيودًا أكثر من اللازم على دالة العرض من خلال استبعادها لأكثر من اللازم من المتغيرات حتى يتسنى توصيف الدالة. هذا الموقف لايعني بالضرورة أن التوصيف أكثر من اللازم يعتبر موقفًا سيئًا حيث سنرى في الفصل (20) كيف يمكن أن نتعامل مع مشكلة وجود معلومات أكثر من اللازم، أو وجود قيود أكثر من اللازم.

والآن لقد تعرضنا إلى كل الحالات الممكنة. وكما اتضح من المناقشة السابقة، فإن أي معادلة في نظام المعادلات الآنية قد تكون موصفة بأقل مما يجب أو موصفة (أكثر مما يجب أو تامة). النموذج ككل يتم توصيفه إذا كانت كل معادلة فيه موصفة. لضمان التوصيف، فإننا نلجأ إلى المعادلات المخفضة. ولكن في فقرة (3.19) سنستعرض طريقة بديلة، وتأخذ وقتًا أقل لتحديد ما إذا كانت معادلة ما في نظام المعادلات الآنية يمكن توصيفها أم لا.

3.19 قواعد التوصيف : RULES FOR IDENTIFICATION

كما يتضح من أمثلة الفقرة 2.19، فمن حيث المبدأ من الممكن اللجوء إلى المعادلات المخفضة لتحديد التوصيف الخاص بمعادلة ما في نظام المعادلات الآنية. ولكن هذه الأمثلة توضح أيضًا كم الوقت والمجهود اللازم لمثل هذه العملية. ولحسن

(4) لاحظ الفرق بين التوصيف بأكثر مما يجب أو بأقل مما يجب. في الحالة الأولى من الممكن الحصول على تقديرات المعاملات البنائية، في حين في الحالة الأخيرة هناك العديد من التقديرات لواحد أو أكثر من المعاملات البنائية.

الحظ ليس من الضروري استخدام هذه الطريقة. الطريقة المسماة بشروط الرتبة والترتيب للتوصيف تجعل هذه المهمة أكثر سهولة، من خلال إتاحة طريقة منتظمة أسهل في الحساب.

لفهم شروط الرتبة والترتيب، دعنا نستعرض الرموز التالية:

M = عدد المتغيرات الداخلية في النموذج

m = عدد المتغيرات الداخلية لمعادلة ما

K = عدد المتغيرات المحددة سابقًا في النموذج بما في ذلك الجزء المقطوع من المحور الصادي

k = عدد المتغيرات المحددة سابقًا لمعادلة ما

الشرط الترتيبي للقدرة على التوصيف:[5]

The Order Condition of Identification

شرط ضروري (ولكنه غير كاف) للتوصيف، ومعروف باسم شرط الترتيب يمكن التعبير عنه بطريقتين مختلفتين، ولكن متكافئتين كالتالي (الشرط الضروري والكافي للتوصيف سيتم استعراضه لاحقًا).

تعريف 1.19

في نموذج M من المعادلات الآنية حتى يمكن القول بأن معادلة ما يمكن توصيفها لابد من استبعاد على الأقل $(M-1)$ من المتغيرات (الداخلية والمحددة سابقًا) الموجودة في النموذج. فإذا تم استبعاد $(M-1)$ متغير بالضبط، فإن المعادلة تكون تامة التوصيف، أما إذا تم استبعاد أكثر من $(M-1)$ متغير، فإن المعادلة تكون موصفة بأكثر مما يجب.

تعريف 2.19

في نموذج M من المعادلات الآنية، حتي يمكن القول بأن معادلة ما يمكن توصيفها، فإن عدد المتغيرات المحددة سابقًا والمستبعدة من المعادلة يجب ألا يكون أقل من عدد المتغيرات الداخلية الموجودة في هذه المعادلة مطروح من 1، بمعنى أن:

$$K-k \geq m-1 \qquad \textbf{(1.3.19)}$$

إذا كانت $K-k=m-1$، فإن المعادلة تكون تامة التوصيف، ولكن إذا كانت $K-k>m-1$ فإن المعادلة موصفة بأكثر مما يجب.

(5) المصطلح ترتيب يشير إلى ترتيب المصفوفة وهو عدد صفوف أو أعمدة المصفوفة. انظر App.B.

في تمرين 1.19، يسأل القارئ عن إثبات أن التعريفين السابقين متكافئين.

لشرح شرط الترتيب، دعنا نعود إلى أمثلتنا السابقة.

مثال 1.19

(1.2.18) دالة الطلب: $Q_t = \alpha_0 + \alpha_1 P_t + u_{1t}$

(2.2.18) دالة العرض: $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$

هذا النموذج يحتوي على متغيرين داخليين P و Q، ولا توجد أي متغيرات محددة سابقًا. حتى يكون النموذج موصفًا فلابد أن تستبعد هذه المعادلات على الأقل $M - 1 = 1$ متغير. وحيث إن هذه ليست الحالة الراهنة، فإن كلاً من المعادلتين لايمكن توصيفهما.

مثال 2.19

(12.2.19) دالة الطلب: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$

(13.2.19) دالة العرض: $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$

في هذا النموذج لدينا P و Q كمتغيرات داخلية، و I متغير خارجي. بتطبيق شرط الترتيب المعطى في (1.3.19) نجد أن دالة الطلب لايمكن توصيفها، في حين أن دالة العرض تكون تامة التوصيف، حيث إنها تستبعد بالضبط $M - 1 = 1$ متغير وهو I_t.

مثال 3.19

(12.2.19) دالة الطلب: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$

(22.2.19) دالة العرض: $Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$

علمًا بأن Q_t، P_t متغيرات داخلي، و I_t و P_t متغيرات محددة سابقًا، فإن المعادلة (12.2.19) تستبعد متغيراً واحداً فقط وهو P_{t-1} وهو المعادلة (22.2.19) تستبعد أيضًا متغيراً واحداً فقط وهو I_t وبالتالي كل من المعادلتين يمكن توصيفهما وفقًا لشرط الترتيب وبالتالي فالنموذج ككل يمكن توصيفه.

مثال 4.19

(28.2.19) دالة الطلب: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t}$

(22.2.19) دالة العرض: $Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$

في هذا النموذج لدينا P_t و Q_t كمتغيرات داخلية و I_t، R_t و P_{t-1} كمتغيرات محددة سابقًا. دالة الطلب تستبعد متغيراً واحداً فقط وهو P_{t-1} وبالتالي باستخدام شرط

الترتيب فإن هذه المعادلة تامة التوصيف. ولكن دالة العرض تستبعد متغيرين اثنين وهما I_t و R_t وبالتالي فإنها موصفة بأكثر مما يجب. وكما لاحظنا من قبل، في مثل هذه الحالة هناك طريقتان لتقدير β_1 والذي يمثل معامل متغير السعر.

لاحظ أن هناك بعض التعقيد في هذا الأمر. وفقًا لشرط الترتيب، فإن دالة الطلب تامة التوصيف، ولكن إذا حاولنا تقدير معلمات هذه المعادلة من المعاملات المخفضة الشكل المعطاة في (31.2.19) فإن التقديرات لن تكون وحيدة، حيث β_1 والتي تدخل ضمن الحسابات ستكون لها قيمتان، وسيكون علينا تحديد أي من هاتين القيمتين أفضل. ولكن تلك الحسابات يمكن تجنبها، حيث إنه كما سنرى في الفصل (20) في حالة التوصيف بأكثر مما يجب، فإن طريقة المربعات الصغرى غير المباشرة تكون غير مناسبة، ويجب عدم استخدامها، واستخدام بعض الطرق الأخرى. إحدى هذه الطرق تسمى طريقة المربعات الصغرى على مرحلتين، والتي ستتم مناقشتها بالتفصيل في الفصل (20).

كما يتضح من الأمثلة السابقة، فإن توصيف معادلة ما في نموذج معادلات آنية، يكون ممكنًا إذا كانت المعادلة تستبعد واحدًا أو أكثر من المتغيرات الموجودة في النموذج. هذا الوضع معروف باسم معيار الاستبعاد (للمتغيرات)، أو معيار القيود الصفرية (معاملات المتغيرات التي لاتظهر في المعادلة، من المفترض أنها تساوي صفرًا) هذا المعيار هو الأكثر استخدامًا لضمان إمكانية توصيف المعادلة. ولكن لاحظ أن معيار القيود الصفرية يعتمد على توقعات مسبقة أو نظرية، على ألا يظهر هذا المتغير (المستبعد) في المعادلة. ويرجع إلى الباحث أن يشرح بوضوح أسباب ظهور أو عدم ظهور المتغير في معادلة ما.

شرط الرتبة للتوصيف:[6] The Rank Condition of Identifiability

شرط الترتيب السابق الحديث عنه، هو شرط ضروري، ولكنه غير كاف للتوصيف، بمعنى أنه حتى إذا توفر هذا الشرط، قد توجد معادلة لايمكن توصيفها، فمثلاً في مثال 2.19، معادلة العرض يمكن توصيفها، حيث إنها وفقًا لشرط الترتيب تستبعد متغير الدخل I_t والذي يظهر في دالة الطلب، ولكن التوصيف ممكن فقط إذا كانت α_2 (معامل I_t) في دالة الطلب لايساوي الصفر، بمعنى أن متغير الدخل ليس فقط محتملاً ولكن بالفعل لايدخل في دالة الطلب.

---

(6) لفظ الرتبة يشير إلى رتبة المصفوفة، وهو عبارة عن أكبر ترتيب للمصفوفة المربعة (في المصفوفة المعطاة) والتي يكون محددها لايساوي الصفر. وبعبارة أخرى رتبة المصفوفة هي أكبر عدد من الصفوف أو الأعمدة المستقلة خطيًا لهذه المصفوفة انظر الملحق .App.B.

بشكل أكثر عمومية حتى إذا تحقق شرط الترتيب $K - k \geq m - 1$ لمعادلة ما قد تكون غير موصفة، حيث إن المتغيرات المحددة سابقًا والمستبعدة من معادلة ما ولكنها موجودة في النموذج قد لاتكون جميعها مستقلة، وبالتالي قد لاتوجد علاقة واحد إلى واحد بين المعاملات البنائية (βs) والمعاملات المخفضة الشكل (Π's)، وبالتالي نكون غير قادرين على تقدير المعاملات البنائية من المعاملات المخفضة، وسنوضح ذلك بالتفصيل لاحقًا، وبالتالي نحن نحتاج إلى شرط ضروري، وفي نفس الوقت كاف للتوصيف، وهذا هو شرط الرتبة للتوصيف والذي يمكن صياغته كالتالي:

شرط الرتبة للتوصيف: في نموذج ما يحتوي على M معادلة في M متغير داخلي. يقال إن معادلة ما يمكن توصيفها إذا وفقط إذا كان على الأقل هناك محدد واحد من الرتبة $(M - 1)(M - 1)$ غير صفري، ومكون من معاملات المتغيرات (الداخلية والمحددة سابقًا معًا) المستبعدة من معادلة ما، ولكن موجودة في معادلات أخرى في النموذج.

لشرح شرط الرتبة للتوصيف، دعنا نستعرض نموذج المعادلات الآنية التالي، الافتراض حيث يوجد Y متغير داخلي ومتغيرات الـ X هي متغيرات محددة سابقًا[7].

$$Y_{1t} - \beta_{10} \qquad - \beta_{12}Y_{2t} - \beta_{13}Y_{3t} - \gamma_{11}X_{1t} = u_{1t} \quad (2.3.19)$$

$$Y_{2t} - \beta_{20} \qquad - \beta_{23}Y_{3t} - \gamma_{21}X_{1t} - \gamma_{22}X_{2t} = u_{2t} \quad (3.3.19)$$

$$Y_{3t} - \beta_{30} - \beta_{31}Y_{1t} \qquad - \gamma_{31}X_{1t} - \gamma_{32}X_{2t} = u_{3t} \quad (4.3.19)$$

$$Y_{4t} - \beta_{40} - \beta_{41}Y_{1t} - \beta_{42}Y_{2t} \qquad - \gamma_{43}X_{3t} = u_{4t} \quad (5.3.19)$$

لتسهيل التوصيف، دعنا نصيغ النظام السابق في جدول (1.19) السهل فهمه. دعنا أولاً نطبق شرط الترتيب للتوصيف، كما موضح في جدول (2.19) وفقًا لشرط الترتيب، فإن كل معادلة قابلة للتوصيف. دعنا الآن نتأكد من شرط الرتبة. فمثلاً بالنسبة للمعادلة الأولى والتي تستبعد المتغيرات Y_4، X_2 و X_3 (ويمثل ذلك بأصفار في الصف الأول في جدول 1.19). حتى تكون هذه المعادلة موصفة، فلابد من أن نحصل على الأقل على محدد غير صفري من الترتيب 3 x 3 من معاملات المتغيرات المستبعدة في هذه المعادلة، ولكنها موجودة في معادلات أخرى. للحصول على

(7) نظام المعادلات الآنية الموجود في (1.1.19) يمكن كتابته في شكل بديل ملائم أكثر للتعامل مع المصفوفات.

المحدد نحصل أولاً على مصفوفة معاملات المتغيرات المناسبة، وهي المتغيرات Y_4، X_2 و X_3 الموجودة في معادلات أخرى. في مثالنا الحالي تكون كالتالي.

جدول (1.19)

Equation no.	Coefficients of the variables							
	1	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3
(19.3.2)	$-\beta_{10}$	1	$-\beta_{12}$	$-\beta_{13}$	0	$-\gamma_{11}$	0	0
(19.3.3)	$-\beta_{20}$	0	1	$-\beta_{23}$	0	$-\gamma_{21}$	$-\gamma_{22}$	0
(19.3.4)	$-\beta_{30}$	$-\beta_{31}$	0	1	0	$-\gamma_{31}$	$-\gamma_{32}$	0
(19.3.5)	$-\beta_{40}$	$-\beta_{41}$	$-\beta_{42}$	0	1	0	0	$-\gamma_{43}$

جدول (2.19)

Equation no.	No. of predetermined variables excluded, $(K-k)$	No. of endogenous variables included less one, $(m-1)$	Identified?
(19.3.2)	2	2	Exactly
(19.3.3)	1	1	Exactly
(19.3.4)	1	1	Exactly
(19.3.5)	2	2	Exactly

مثل هذه المصفوفة، دعنا نسميها **A**، تكون كالتالي:

$$\mathbf{A} = \begin{bmatrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{bmatrix} \tag{6.3.19}$$

من الممكن ملاحظة أن محدد هذه المصفوفة يساوي الصفر:

$$\det \mathbf{A} = \begin{vmatrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{vmatrix} \tag{7.3.19}$$

وبما أن المحدد يساوي الصفر، فإن رتبة المصفوفة (6.3.19) والتي نرمز لها بالرمز $\rho(\mathbf{A})$ أقل من 3. وبالتالي المعادلة (2.3.19) لاتحقق شرط الرتبة، وبالتالي لايمكن توصيفها.

كما لاحظنا فإن شرط الرتبة يعتبر ضروريًا وكافيًا لتحقيق التوصيف. وبالتالي على الرغم من أن شرط الترتيب يوضح أن المعادلة (2.3.19) يمكن توصيفها، إلا أن شرط الرتبة ينفي ذلك. ونستطيع أن نرى أن أعمدة أو صفوف المصفوفة **A** المعطاة في (6.3.19) ليست مستقلة (خطيًا)، بمعنى أن هناك بعض العلاقات بين المتغيرات Y_4، X_2 و X_3. وكنتيجة لذلك، لن تكون لدينا معلومات كافية لتقدير معلمات المعادلة

(2.3.19)، المعادلات المخفضة للنموذج السابق ستوضح أنه من غير الممكن الحصول على المعاملات البنائية من المعاملات المخفضة الشكل. سيطلب من القارئ إثبات أن المعادلات (3.3.19) و (4.3.19) غير قابلة للتوصيف وفقًا لشرط الرتبة ولكن المعادلة (5.3.19) يمكن توصيفها.

كما أوضحنا في المناقشة السابقة، شرط الرتبة يحدد لنا ما إذا كانت المعادلة محل الدراسة يمكن توصيفها أم لا، في حين أن شرط الترتيب يحدد ما إذا كانت تامة التوصيف أو مصفوفة بأكثر مما يجب.

لتطبيق شرط الرتبة يمكن اتباع التالي:

1 - اكتب النظام محل الدراسة في شكل جدول، كما هو موضح في جدول (1.19).

2 - قم بعملية ضرب للمعاملات الموجودة في صف المعادلة محل الدراسة.

3 - أيضًا اضرب الأعمدة المرتبطة بهذه المعاملات في 2 والتي تكون معاملات غير صفرية.

4 - الجانب المتبقي من الجدول سيعطي فقط معاملات المتغيرات الموجودة في النظام، ولكن ليست في المعادلة محل الدراسة. من كل هذه المدخلات كون كل المصفوفات الممكنة، مثل **A**، من الرتبة $M-1$ واحصل على المحدد الخاص بكل مصفوفة. إذا كان على الأقل هناك محدد واحد غير صفري، فإن المعادلة محل الدراسة يمكن توصيفها (تامة أو مصفوفة أكثر مما يجب). رتبة المصفوفة، مثلاً **A**، في مثل هذه الحالة، تكون مساوية تمامًا لـ $M-1$. إذا كانت كل الـ $(M-1)(M-1)$ محدد يساوي الصفر، فإن رتبة المصفوفة **A** تكون أقل من $M-1$ والمعادلة محل الدراسة تكون غير مصفوفة.

مناقشتنا بخصوص شروط الرتبة والترتيب للتوصيف، تحدد لنا المبادئ العامة التالية لتوصيف أي معادلة بنائية في نظام ما مكون من M من المعادلات الآنية:

1- إذا كان $K-k>m-1$ ورتبة المصفوفة **A** هي $M-1$ تكون المعادلة موصفة بأكثر مما يجب.
2- إذا كان $K-k=m-1$ ورتبة المصفوفة **A** هي $M-1$ فإن المعادلة تكون تامة التوصيف.
3- إذا كان $K-k\geq m-1$ ورتبة المصفوفة **A** أقل من $M-1$ فإن المعادلة موصفة بأقل مما يجب.
4- إذا كان $K-k>m-1$ فإن المعادلة البنائية لايمكن توصيفها ورتبة المصفوفة **A** في مثل هذه الحالة تكون محددة بأقل من $M-1$ (لماذا؟).

وبالتالي، عندما نتحدث عن التوصيف، فإننا نعني إما التوصيف التام أو التوصيف بأكثر مما يجب. لايوجد داع لاعتبار حالة عدم التوصيف أو التوصيف بأقل مما يجب، حيث إنه بغض النظر عن مدى شمول البيانات، فإن المعلمات البنائية لايمكن تقديرها. عمومًا كما سنرى في الفصل (20)، فإن معاملات المعادلات تامة التوصيف أو الموصفة أكثر مما يجب يمكن تقديرها. السؤال الآن أي من شرطي الترتيب أو الرتبة يمكن استخدامه في الجانب العملي؟ في نماذج المعادلات الآنية الكبيرة، تطبيق شرط الرتبة يعتبر مهمة صعبة، وبالتالي كما ذكر Harvey فإن: لحسن الحظ، فإن شرط الترتيب عادة كاف لضمان التوصيف، وعلى الرغم من أهمية الدراية بشرط الرتبة، ولكن صعوبة تطبيقه ستؤدي أحيانًا إلى كارثة[8].

4.19[*] اختبار الآنية :[9] A TEST OF SIMULTANEITY

إذا لم توجد معادلة آنية أو مشكلة الآنية، فإن مقدرات الـ OLS تعطي تقديرات متسقة وكافية. على الجانب الآخر. إذا كانت هناك مشكلة الآنية، فإن مقدرات الـ OLS ليست حتى متسقة. في وجود الآنية، كما سنرى في الفصل (20)، طريقة المربعات الصغرى على مرحلتين (2SLS) والمتغيرات المساهمة ستعطي تقديرات متسقة وكافية. وعلى العكس، إذا طبقنا هذه الطرق في حالة عدم وجود حقيقي لمشكلة الآنية، فإن هذه الطرق ستعطي مقدرات متسقة ولكن غير كافية (بمعنى تباين أقل) كل هذا يستدعي اختبار وجود مشكلة الآنية قبل استبعاد استخدام الـ OLS وتفضيل استخدام الطرق البديلة الأخرى.

كما أوضحنا من قبل، مشكلة الآنية تظهر عندما تكون هناك بعض المتغيرات المنحدرة كمتغيرات داخلية، وبالتالي محتمل أن تكون مرتبطة مع مقدار الخطأ. وبالتالي من الضروري وجود اختبار للآنية لمعرفة ما إذا كان المتغير المنحدر (الداخلي) مرتبطًا مع مقدار الخطأ أم لا. إذا كان هناك ارتباط، فإن مشكلة الآنية تكون موجودة، وبالتالي لابد من استخدام الطرق البديلة لطريقة الـ OLS، ولكن إذا لم توجد هذه المشكلة، فإننا نستطيع استخدام الـ OLS.

لتحديد ذلك في وضع واقعي، دعنا نستخدم اختبار Hausman لتحديد الخطأ.

(8) Andrew Harvey, The Econometric Analysis of Time Series, 2d ed., The MIT Press, Cambridge, Mass., 1990, p.328.

(*) اختياري

(9) المناقشة التالية مأخوذة من Robert and Daniel L. Rubinfeld, Econometric Models and Economic Pindyck 3. Forecasts, 3d ed., McGrqw-Hill, New York, 1991, pp. 303-305

اختبار Hausman للتحديد: Hausman Specification Test

أحد أشكال اختبار Hausman لتحديد الخطأ، والذي يمكن استخدامه في حالة وجود مشكلة الآنية، يمكن وصفه كالتالي[10]. لتحديد الأفكار، دعنا نعتبر النموذج ذا المعادلتين التالي:

(1.4.19) معادلة الطلب: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t}$

(2.4.19) معادلة العرض: $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$

بحيث إن: P = السعر

Q = الكمية

I = الدخل

R = الثروة

u's = مقادير الأخطاء

افترض أن كلاً من I و R متغيرات خارجية. وبالطبع P و Q متغيرات داخلية. الآن اعتبر دالة العرض (2.4.19). إذا لم توجد مشكلة الآنية (بمعنى أن P و Q مستقلان بالتبعية)، فإن P_t و u_{2t} يكونان غير مرتبطين (لماذا؟). على الجانب الآخر إذا وجدت مشكلة الآنية فإن P_t و u_{2t} سيكونان مرتبطين. لمعرفة أي من هاتين الحالتين موجود، فإن اختبار Hausman يحدث كالتالي:

أولاً من المعادلة (1.4.19) و (2.4.19) نحصل على المعادلات المخفضة الشكل التالية:

$$P_t = \Pi_0 + \Pi_1 I_t + \Pi_2 R_t + v_t \qquad (3.4.19)$$

$$Q_t = \Pi_3 + \Pi_4 I_t + \Pi_5 R_t + w_t \qquad (4.4.19)$$

بحيث إن v و w مقادير أخطاء مخفضة الشكل، تقدير (3.4.19) بالـ OLS نحصل على:

$$\hat{P}_t = \hat{\Pi}_0 + \hat{\Pi}_1 I_t + \hat{\Pi}_2 R_t \qquad (5.4.19)$$

(10) J. A. Hausan, "Specification Tests in Econometrics, "Econometrica, vol. 46, November 1976, pp. 1251–1271. See also A. Nakamura and M. Nakamura, "On the Relationship amog Several Specification Error Tests Presented by Durbin, Wu, and Hausman," Econometrica, vol. 49. November 1981, pp. 1583–1588.

وبالتالي :

$$P_t = \hat{P}_t + \hat{v}_t \quad (6.4.19)$$

بحيث إن $\hat{P}_t$ هي تقدير P_t و $\hat{v}_t$ هو تقدير البواقي. بالتعويض عن (6.4.19) في (2.4.19) نحصل على :

$$Q_t = \beta_0 + \beta_1 \hat{P}_t + \beta_1 \hat{v}_t + u_{2t} \quad (7.4.19)$$

لاحظ أن: معاملات P_t و v_t متساوية.

الآن تحت صحة الفرض العدمي القائل أنه لاتوجد آنية، فإن الارتباط بين $\hat{v}_t$ و u_{2t} لابد أن يساوي الصفر تقاربيًا. وبالتالي إذا أجرينا الانحدار الموجود في المعادلة (7.4.19) ووجدنا أن معامل v_t في (7.4.19) إحصائيًا يساوي الصفر، نستطيع استنتاج عدم وجود مشكلة الآنية وبالطبع هذا الاستنتاج سيكون بالعكس إذا وجدنا أن هذا المعامل له معنوية إحصائية.

وبالتالي فإن اختبار Hausman يشتمل على الخطوات التالية :

الخطوة 1: قم بانحدار P_t على I_t و R_t واحصل على $\hat{v}_t$.

الخطوة 2: قم بإنحدار Q_t على $\hat{P}_t$ و $\hat{v}_t$ وقم بعمل اختبار t لمعامل $\hat{v}_t$ إذا كان المعامل معنويًا، لاترفض الفرض العدمي للآنية وبخلاف ذلك ارفض الفرض العدمي[11]. للحصول على تقدير كاف، فإن Pindyck و Rubinfeld اقترحا انحدار Q_t على P_t و $\hat{v}_t$ [12].

مثال 5.19

نموذج Pindyct- Rubinfeld للصرف العام[13]

لدراسة سلوك إحدى الولايات الأمريكية، والصرف الحكومي المحلي، قام الكتاب باقتراح نموذج المعادلات الآنية التالي :

$$\text{EXP} = \beta_1 + \beta_2\text{AID} + \beta_3\text{INC} + \beta_4\text{POP} + u_i \quad (8.4.19)$$

$$\text{AID} = \delta_1 + \delta_2\text{EXP} + \delta_3\text{PS} + v_i \quad (9.4.19)$$

(11) إذا كان هناك أكثر من متغير منحدر داخلي، لابد أن نستخدم اختبار F.
Pindyc Kand Rubin feld, op.cit., p.3.4.

(12) لاحظ أن المتغير المنحدر هو P_t وليس $\hat{P}_t$ المتغير.

(13) Pindyc Kand Rubin feld, op.cit.,pp. 179-177. notations slights altered.

بحيث إن:

EXP = المصاريف العامة الحكومية المحلية والخاصة بالولاية.

AID = مستوى المعونات الفيدرالية للمساعدات.

INC = دخل الولاية.

POP = عدد سكان الولاية.

PS = عدد الأطفال في المدارس الإعدادية والثانوية.

u, v = مقادير الأخطاء.

في هذا النموذج INC، POP و PS هي متغيرات داخلية.

بسبب الآنية المحتملة بين EXP و AID قام الباحثان أولاً بانحدار AID على INC، POP وPS (بمعنى انحدار مخفض الشكل). افترض أن مقدار الخطأ في هذا الانحدار هو w_i. من هذا الانحدار فإن البواقي المحسوبة هي $\hat{w}_i$، الباحثان قاما بعد ذلك بانحدار EXP على AID، INC، POP و $\hat{w}_i$ وحصلا على النتائج التالية:

$$\widehat{EXP} = -89.41 + 4.50AID + 0.00013INC - 0.518POP - 1.39\hat{w}_i$$

$$t = (-1.04) \quad (5.89) \quad (3.06) \quad (-4.63) \quad (-1.73) \qquad (10.4.19)^{(14)}$$

$$R^2 = 0.99$$

عند مستوى معنوية 5%، معامل $\hat{w}_i$ ليس له معنوية إحصائية، وبالتالي عند هذا المستوى لاتوجد مشكلة الآنية. عموماً عند مستوى معنوية 10% يكون هذا المعامل معنوياً وتظهر مشكلة الآنية.

بشكل عارض فإن مقدرات OLS لـ (8.4.19) هي كالتالي:

$$\widehat{EXP} = -46.81 + 3.24AID + 0.00019INC - 0.597POP$$

$$t = (-0.56) \quad (13.64) \quad (8.12) \quad (-5.71) \qquad (11.4.19)$$

$$R^2 = 0.993$$

لاحظ خاصية مثيرة في النتائج المعطاة في (10.4.19) و (11.4.19): عند وضع مشكلة الآنية في الاعتبار، فإن المتغير AID يصبح أقل معنوية على الرغم من الزيادة الرقمية في قيمته.

5.19 اختبارات لخارجية النشأة:(*) TESTS FOR EXOGENEITY

سبق وأن ذكرنا من قبل، أن من مسئوليات الباحث تحديد أي من المتغيرات محل الدراسة سيعتبرها متغيرات داخلية أو خارجية. وذلك سيعتمد على المشكلة محل البحث، وأي معلومات سابقة موجودة لدى الباحث عن الموضوع. ولكن هل من الممكن إيجاد اختبار إحصائي لخارجية المتغيرات، على غرار اختبار Granger للسببية؟

(14) في الملاحظة 12، الباحثان استخدما AID بدلاً من $\widehat{AID}$ كمتغير منحدر.

(*) اختياري

اختبار Hausman الذي تمت مناقشته في الفقرة 4.19 يقدم إجابة لهذا السؤال. افترض أن لدينا نموذجًا به ثلاث معادلات في ثلاثة متغيرات داخلية Y_1، Y_2 و Y_3 وافترض وجود ثلاثة متغيرات خارجية X_1، X_2 و X_3 وافترض أن المعادلة الأولى في النموذج هي كالتالي:

$$Y_{1i} = \beta_0 + \beta_2 Y_{2i} + \beta_3 Y_{3i} + \alpha_1 X_{1i} + u_{1i} \qquad \textbf{(1.5.19)}$$

إذا كان Y_2 و Y_3 متغيرات داخلية حقيقية، لا تستطيع تقدير (1.5.19) باستخدام OLS (لماذا؟). ولكن كيف تستطيع معرفة ذلك؟ يمكن القيام بالتالي. نحصل على المعادلات المخفضة لكل من Y_2 و Y_3 (لاحظ أن: المعادلات المخفضة الشكل سيكون فيها متغيرات سابقة التحديد فقط على الجانب الأيمن). من هذه المعادلات المخفضة نحصل على $\hat{Y}_{2i}$ و $\hat{Y}_{3i}$ والتي تمثل القيم المتنبأ بها لكل من Y_{2i} و Y_{3i} بالترتيب. وبنفس الفكرة اختبار Hausman المناقش سابقًا، فإننا نستطيع تقدير المعادلة التالية باستخدام OLS كالتالي:

$$Y_{1i} = \beta_0 + \beta_2 Y_{2i} + \beta_3 Y_{3i} + \alpha_1 X_{1i} + \lambda_2 \hat{Y}_{2i} + \lambda_3 \hat{Y}_{3i} + u_{1i} \qquad \textbf{(2.5.19)}$$

باستخدام اختبار F، نستطيع اختبار الفرض: $\lambda_2 = \lambda_3 = 0$. إذا رفضنا هذا الفرض، فإن Y_2 و Y_3 يكونا متغيرين داخليين ولكن إذا لم نرفض هذا الفرض، يمكن التعامل معهما على أنهما متغيران خارجيان. لمثال تطبيقي، انظر تمرين 19.16.

6.19 الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

1 – مشكلة التوصيف تسبق مشكلة التقدير.

2 – مشكلة التوصيف تتعلق بإمكانية الحصول على قيم رقمية وحيدة للمعاملات البنائية من المعاملات المخفضة الشكل المقدرة.

3 – إذا كان هذا ممكن، فإنه يقال عن هذه المعادلة داخل نظام المعادلات الآنية، إنها معادلة يمكن توصيفها، أما إذا كان ذلك غير ممكن، فإن المعادلة تكون غير موصفة أو موصفة بأقل مما يجب.

4 – المعادلة الموصفة ممكن أن تكون تامة التوصيف أو موصفة بأكثر مما يجب، في الحالة الأولى تكون هناك قيم وحيدة المعاملات البنائية، ولكن في الحالة الثانية يمكن أن يوجد أكثر من قيمة واحدة لواحد أو أكثر من المعلمات البنائية.

5 - مشكلة التوصيف تظهر بسبب أن نفس البيانات قد تكون مناسبة مع مجموعات مختلفة من المعاملات البنائية، أي مجموعات مختلفة من النماذج. فمثلاً في انحدار السعر على الكمية فقط يكون من الصعب معرفة ما إذا كنا نقدر دالة عرض أو نقدر دالة طلب، حيث إن السعر والكمية موجودان في كلتا الدالتين.

6 - لمعرفة إمكانية توصيف معادلة بنائية ما، يمكن تطبيق أسلوب المعادلات المخفضة الشكل، والتي يتم فيها التعبير عن المتغير الداخلي بمفرده كدالة في متغيرات مسبقة التحديد.

7 - عمومًا، فإن هذه الطريقة والتي تستهلك وقتًا طويلاً يمكن تجنبها باللجوء إلى شرط الترتيب أو شرط الرتبة للتوصيف. وعلى الرغم من أن شرط الترتيب يسهل تطبيقه، فإنه يعتبر شرطًا ضروريًا ولكن غير كاف للتوصيف، ولكن على الجانب الآخر، فإن شرط الرتبة يعتبر شرطًا ضروريًا وكافًا للتوصيف. فإذا تحقق شرط الرتبة، فإن هذا يعني أن شرط الترتيب متحقق أيضًا ولكن العكس غير صحيح. في الواقع العملي، فإن شرط الترتيب عمومًا يضمن تحققه إمكانية التوصيف.

8 - في وجود مشكلة الآنية يكون من غير الممكن تطبيق الـ OLS كما سبق وذكرنا في الفصل (18). ولكن إذا رغب الباحث في استخدامها، فلابد من اختبار الآنية بشكل واضح. اختبار Hausman للتحديد يمكن أن يستخدم لهذا الغرض.

9 - على الرغم من أن تحديد ما إذا كان المتغير متغيراً داخليًا أو خارجيًا يرجع إلى حكم الباحث، إلا أنه من الممكن استخدام اختبار Hausman لتحديد ما إذا كان متغير ما أو مجموعة من المتغيرات داخلية أو خارجية.

10 - على الرغم من انتماء مفهوم السببية وخارجية المنشأ إلى نفس العائلة، إلا أنهما مختلفان، وقد لايعني أحدهما بالضرورة الآخر. في الواقع العملي من الأفضل الفصل بين هذين المفهومين (انظر الفقرة 14.17).

تمـــاريـــن : EXERCISES

1.19 اثبت أن التعريفين الخاصين بشرط الرتبة للتوصيف متساويان.

2.19 استنتج المعاملات البنائية من المعاملات المخفضة المعطاة في (25.2.19) و(27.2.19).

3.19 احصل على الشكل المخفض للنماذج التالية، وحدد في كل حالة ما إذا كانت المعادلات البنائية غير موصفة، تامة التوصيف أو موصفة بأكثر مما يجب.

(a) الفصل 18، مثال 2.18.

(b) الفصل 18، مثال 3.18.

(c) الفصل 18، مثال 6.18.

4.19 تحقق من إمكانية توصيف النماذج الموجودة في تمرين 3.19 باستخدام كل من شرط الترتيب، وشرط الرتبة للتوصيف.

5.19 في النموذج (22.2.19) و (28.2.19) تم توضيح أن معادلة العرض موصفة بأكثر مما يجب. ما هي القيود، إن وجدت اللازم فرضها على المعلمات البنائية التي تجعل هذه المعادلة تامة التوصيف؟ علل اختيارك لهذه القيود.

6.19 من النموذج

$$Y_{1t} = \beta_{10} + \beta_{12} Y_{2t} + \gamma_{11} X_{1t} + u_{1t}$$

$$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} + \gamma_{22} X_{2t} + u_{2t}$$

تم الحصول على المعادلات المخفضة التالية :

$$Y_{1t} = \Pi_{10} + \Pi_{11} X_{1t} + \Pi_{12} X_{2t} + w_t$$

$$Y_{2t} = \Pi_{20} + \Pi_{21} X_{1t} + \Pi_{22} X_{2t} + v_t$$

(a) هل المعادلات البنائية يمكن توصيفها؟

(b) ماذا يحدث للتوصيف إذا كان معلومًا من قبل أن $\gamma_{11} = 0$؟

7.19 بالعودة إلى تمرين 6.19 المعادلات المخفضة المقدرة هي كالتالي:

$$Y_{1t} = 4 + 3X_{1t} + 8X_{2t}$$

$$Y_{2t} = 2 + 6X_{1t} + 10X_{2t}$$

(a) احصل على قيم المعلمات البنائية.

(b) كيف يمكن اختبار الفرض العدمي: $\gamma_{11} = 0$؟

$$Y_{1t} = \beta_{10} + \beta_{12} Y_{2t} + \gamma_{11} X_{1t} + u_{1t}$$
$$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} + u_{2t}$$

8.19 اعتبر النموذج التالي:

$$Y_{1t} = 4 + 8X_{1t}$$
$$Y_{2t} = 2 + 12X_{1t}$$

وحصلنا منه على المعادلات المخفضة التالية:

(a) أي من المعاملات البنائية، إذا وجد، يمكن تقديره من المعاملات المخفضة؟ علل استنتاجك.

(b) كيف ستتغير الإجابة على (a) إذا علم مسبقًا أن: (1) $\beta_{12}=0$ و (2) $\beta_{10}=0$

9.19 حدد ما إذا كانت المعادلات البنائية في النموذج المعطى في تمرين 8.18 موصفة أم لا.

10.19 بالعودة إلى تمرين 7.18 حدد أيًا من المعادلات البنائية يمكن توصيفها.

11.19 جدول (3.19) يمثل نموذجًا في خمس معادلات بخمسة متغيرات داخلية Y وأربعة متغيرات خارجية X.

جدول (3.19)

Equation no.	Coefficients of the variables								
	Y_1	Y_2	Y_3	Y_4	Y_5	X_1	X_2	X_3	X_4
1	1	β_{12}	0	β_{14}	0	γ_{11}	0	0	γ_{14}
2	0	1	β_{23}	β_{24}	0	0	γ_{22}	γ_{23}	0
3	β_{31}	0	1	β_{34}	β_{35}	0	0	γ_{33}	γ_{34}
4	0	β_{42}	0	1	0	γ_{41}	0	γ_{43}	0
5	β_{51}	0	0	β_{54}	1	0	γ_{52}	γ_{53}	0

حدد إمكانية كل معادلة باستخدام شرط الترتيب، وشرط الرتبة للتوصيف.

12.19 اعتبر نموذج Keynesian الموسع لتحديد الدخل التالي:

دالة الاستهلاك: $C_t = \beta_1 + \beta_2 Y_t - \beta_3 T_t + u_{1t}$

دالة الاستثمار: $I_t = \alpha_0 + \alpha_1 Y_{t-1} + u_{2t}$

دالة الضرائب: $T_t = \gamma_0 + \gamma_1 Y_t + u_{3t}$

وحدة الدخل: $Y_t = C_t + I_t + G_t$

حيث إن:

C = مصاريف الاستهلاك

Y = الدخل

I = الاستثمار

T = الضرائب

G = مصاريف حكومية

u's = مقادير الأخطاء

في النموذج المتغيرات الداخلية هي C، I، T و Y والمتغيرات المحددة مسبقًا هي G و Y_{t-1}.

بتطبيق شرط الترتيب، تأكد من إمكانية توصيف كل معادلة داخل النظام، وإمكانية توصيف النظام ككل. ماذا سيحدث إذا افترض أن r_t، معدل الفائدة، متغير خارجي يظهر في الجانب الأيمن من معادلة الاستثمار؟

13.19 بالعودة إلى البيانات المعطاة في جدول (1.18) في الفصل (18). باستخدام هذه البيانات، قدر الانحدار المخفض (2.1.19) و (4.1.19). هل يمكنك تقدير β_0 وβ_1؟ وضح خطواتك الحسابية. هل النموذج يمكن توصيفه؟ علل إجابتك.

14.19 افترض التعريف كالتالي كتعريف آخر بشرط الترتيب للتوصيف:

$$K \geq m + k - 1$$

والذي ينص على أن عدد المتغيرات المحددة سابقًا في النموذج لايمكن أن يكون أقل من عدد المعاملات المجهولة في المعادلة المراد توصيفها. اثبت أن هذا التعريف مماثل للتعريفين الآخرين اللذين تم عرضهما في فقرة شرط الترتيب للتوصيف.

15.19 التالي هو نسخة مبسطة من نموذج Suits لسوق البطيخ(*):

دالة الطلب:

$$P_t = \alpha_0 + \alpha_1(Q_t/N_t) + \alpha_2(Y_t/N_t) + \alpha_3 F_t + u_{1t}$$

دالة عرض المحصول:

$$Q_t = \beta_0 + \beta_1(P_t/W_t) + \beta_2 P_{t-1} + \beta_3 C_{t-1} + \beta_4 T_{t-1} + u_{2t}$$

بحيث إن: P = السعر

(Q/N) = الكمية المطلوبة لكل فرد

(*) D.B. Suits, "An Econometric Model of the Watermelon Market," Journal of Farm Economics, vol. 37, 1955, pp.237- 251.

(Y/N) = الدخل الفردي

F_t = تكلفة التبريد

(P/W) = السعر بالنسبة إلى معدل الأجر في المزرعة

C = سعر القطن

T = سعر الخضراوات الأخرى

N = عدد السكان

P ، Q تعتبر متغيرات داخلية.

(**a**) احصل على الشكل المخفض.

(**b**) حدد ما إذا كانت دالة العرض، دالة الطلب أو كلاهما يمكن توصيفها.

16.19 اعتبر نموذج الطلب والعرض على المال التالي:

طلب المال : $M_t^d = \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 P_t + u_{1t}$

عرض المال : $M_t^s = \alpha_0 + \alpha_1 Y_t + u_{2t}$

بحيث إن: M = المال

Y = الدخل

R = معدل الفائدة

P = السعر

u's = مقادير الأخطاء

جدول (14.19) المال، GDP، معدل الفائدة ومؤشر سعر المستهلك، الولايات المتحدة 1970–1999

Observation	M_2	GDP	TBRATE	CPI
1970	626.4000	3578.000	6.458000	38.80000
1971	710.1000	3697.700	4.348000	40.50000
1972	802.1000	3998.400	4.071000	41.80000
1973	855.2000	4123.400	7.041000	44.40000
1974	901.9000	4099.000	7.886000	49.30000
1975	1015.900	4084.400	5.838000	53.80000
1976	1151.700	4311.700	4.989000	56.90000
1977	1269.900	4511.800	5.265000	60.60000
1978	1365.500	4760.600	7.221000	65.20000
1979	1473.100	4912.100	10.04100	72.60000
1980	1599.100	4900.900	11.50600	82.40000
1981	1754.600	5021.000	14.02900	90.90000
1982	1909.500	4913.300	10.68600	96.50000
1983	2126.000	5132.300	8.630000	99.60000
1984	2309.700	5505.200	9.580000	103.9000
1985	2495.400	5717.100	7.480000	107.6000
1986	2732.100	5912.400	5.980000	109.6000

1987	2831.100	6113.300	5.820000	113.6000
1988	2994.300	6368.400	6.690000	118.3000
1989	3158.400	6591.900	8.100000	124.0000
1990	3277.600	6707.900	7.510000	130.7000
1991	3376.800	6676.400	5.420000	136.2000
1992	3430.700	6880.000	3.450000	140.3000
1993	3484.400	7062.600	3.020000	144.5000
1994	3499.000	7347.700	4.290000	148.2000
1995	3641.900	7543.800	5.510000	142.4000
1996	3813.300	7813.200	5.020000	156.9000
1997	4028.900	8159.500	5.070000	160.5000
1998	4380.600	8515.700	4.810000	163.0000
1999	4643.700	8875.800	4.660000	166.6000

لاحظ أن: M_2 = المعروض من المال M_2 (بليون دولار، معدلة موسميًا)

GDP = إجمالي الناتج المحلي (بليون دولار، معدلة موسميًا).

TBRATE = معدل وزارة المالية للورقة النقدية كل 3 أشهر، %

CPI = مؤشر سعر المستهلك (1982–1984 = 100)

المصدر: Economic Report of the president 2001, tables B-2, B-60m B-73, B-69

افترض أن R و P متغيرات خارجية و M و Y متغيرات داخلية. جدول (4.19) يعطي بيانات عن M (معرفة M_2)، Y (GDP)، R (معدل وزارة المالية للورقة المالية [3 أشهر])، و P (مؤشر سعر المستهلك) للولايات المتحدة في الفترة 1970–1999.

(a) هل معادلة الطلب يمكن توصيفها؟

(b) هل معادلة العرض يمكن توصيفها؟

(c) احصل على المعادلات المخفضة الشكل لكل من M و Y.

(d) طبق اختبار الآنية على دالة العرض.

(e) كيف يمكنك معرفة ما إذا كان Y في دالة عرض المال في الحقيقة متغير داخلي؟

الفصل العشرون

طرق المعادلات الآنية

SIMULTANEOUS- EQUATION METHODS

في الفصلين السابقين، قمنا بمناقشة طبيعة نماذج المعادلات الآنية، في هذا الفصل، سنستعرض مشكلة تقدير معالم تلك النماذج. كبداية، يمكن ملاحظة أن مشكلة التقدير أكثر تعقيداً، حيث يوجد العديد من أساليب التقدير المختلفة ذات الخصائص الإحصائية المتعددة. وبالنظر إلى طبيعة الموضوع محل البحث، سنقوم بدراسة عدد محدود من أساليب التقدير المختلفة. دراستنا ستكون بسيطة، وغالباً ستساعد القارئ على المزيد من المعرفة، ولكن التفاصيل العميقة للموضوع متروك للقارئ متابعتها في المراجع المذكورة.

1.20 أساليب التقدير : APPROACHES TO ESTIMATION

دعنا نعتبر نموذجاً عاماً يحتوي على M معادلة بها M من المتغيرات الداخلية معطاة في (1.1.19)، سنستعرض طريقتين لتقدير المعادلات البنائية، وهما طرق المعلومات المحدودة والطريقة الأخرى هي طرق النظام، والمعروفة أيضاً باسم طرق المعلومات الكاملة. في طرق المعادلة الفردية، نقدر كل معادلة في النظام (نظام من المعادلات الآنية) بشكل منفرد آخذين في الاعتبار أي قيود مفروضة على هذه المعادلة (مثل استبعاد بعض المتغيرات)، بغض النظر عن القيود المفروضة على المعادلات الأخرى في النظام[1]، ونظراً لذلك، سميت هذه الطريقة باسم طريقة المعلومات المحدودة. على الجانب الآخر، في طرق النظام نقدر كل المعادلات في النموذج آنياً،

(1) من أجل التوصيف يتم عموماً أخذ معلومات المعادلات الأخرى في الاعتبار. لكن كما لاحظنا في الفصل (19)، التقدير يكون ممكناً فقط في حالة ما تكون المعادلة تامة التوصيف أو موصفة بأكثر مما يجب. في هذا الفصل، سنفترض أن مشكلة التوصيف تم حلها باستخدام الأساليب التي تم استعراضها في الفصل (19).

آخذين في الاعتبار كل القيود على هذه المعادلات عن طريق حذف أو إلغاء لبعض المتغيرات (تذكر أنه للتوصيف في هذه القيود تعتبر أشياء رئيسية) ومن هنا سميت طرق النظام بطرق المعلومات الكاملة.

كمثال، دعنا نستعرض النموذج التالي، والذي يحتوى على الأربع معادلات التالية:

$$
\begin{aligned}
Y_{1t} &= \beta_{10} + \beta_{12}Y_{2t} + \beta_{13}Y_{3t} + \gamma_{11}X_{1t} + u_{1t} \\
Y_{2t} &= \beta_{20} + \beta_{23}Y_{3t} + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + u_{2t} \\
Y_{3t} &= \beta_{30} + \beta_{31}Y_{1t} + \beta_{34}Y_{4t} + \gamma_{31}X_{1t} + \gamma_{32}X_{2t} + u_{3t} \\
Y_{4t} &= \beta_{40} + \beta_{42}Y_{2t} + \gamma_{43}X_{3t} + u_{4t}
\end{aligned}
\qquad (1.1.20)
$$

حيث إن الـ Y's هي المتغيرات الداخلية، والـ X's هي المتغيرات الخارجية. إذا رغبنا في التقدير، مثلاً، المعادلة الثالثة، فباستخدام طرق المعادلة المنفردة سنعتبر هذه المعادلة فقط مع استبعاد المتغيرات Y_2 و X_3 منها. على الجانب الآخر في طرق الأنظمة سنحاول تقدير كل المعادلات الأربع آنيًا آخذين في الاعتبار كل القيود المفروضة على كل المعادلات المختلفة الموجودة في النظام.

للاحتفاظ بطبيعة نماذج المعادلات الآنية، فمن الأفضل أن يتم استخدام طرق الأنظمة مثل طريقة الإمكان الأعظم ذات المعلومات الكاملة (FIML)[2]. في الواقع العملي عمومًا مثل هذه الطرق غير شائعة الاستخدام، لأسباب عديدة.

أولاً: تحتوي هذه الطريقة على العديد من العمليات الحسابية المعقدة فمثلاً نموذج Klein- Goldneryer لسنة 1955 والذي يحتوي على عدد صغير نسبيًا من المعادلات (20 معادلة) وخاص بالاقتصاد الأمريكي يوجد به 151 معاملاً غير صفري، وقام الباحث بتقدير 51 معاملاً فقط باستخدام بيانات السلاسل الزمنية. نموذج الاقتصاد القياسي لمركز أبحاث العلوم الاجتماعية Brookings (SSRC) والخاص بالاقتصاد الأمريكي والمنشور في سنة 1965 يشتمل على 150 معادلة[3]. وعلى الرغم من أن مثل هذه النماذج أضافت تفاصيل مهمة لقطاعات عديدة داخل الاقتصاد، فإن إجراء

---

(2) لمناقشة أبسط لهذا الموضوع/ انظر في:

Carl F. Christ, Econometric, Models and Methods John Wiley & Sons, New York, 1966, pp.395- 401.

(3) James S. Duesenberry, Gary Fromm, Lawrence R. Klein, and Edwin Kuh, eds., A Quarterly Model of the United States Economry, Rand McNally, Chicago, 1965.

العمليات الحسابية المرتبطة بها تعتبر مهمة شديدة الصعوبة، حتى مع التقدم التكنولوجي الحالي، بالإضافة على التكلفة المرتفعة الخاصة بها.

ثانيًا: طرق الأنظمة مثل FIML، تؤدي إلى نتائج غير خطية تمامًا في المعلمات، وبالتالي يكون من الصعب تحديدها. ثالثًا: إذا كان هناك خطأ في التوصيف (مثل شكل دالي غير صحيح أو استبعاد لمتغير مهم) في معادلة أو أكثر داخل النظام، فإن هذا الخطأ سينتقل إلى باقي النظام. وكنتيجة لذلك، فإن طرق الأنظمة شديدة الحساسية لأخطاء التعريف (التوصيف).

وبالتالي، فالواقع العملي يتم استخدام طرق المعادلة الفردية بشكل أكثر، وكما قال Klein.

طرق المعادلة الفردية، في مجال الأنظمة الآنية، أقل حساسية لأخطاء التعريف، بمعنى أن أجزاء النظام التي تم تعريفها جيدًا قد لاتتأثر بشكل كبير بأخطاء التعريف الموجودة في الأجزاء الأخرى من النظام[4].

في باقي الفصل، سنتعامل مع طرق المعادلة الفردية فقط، وبشكل محدد، سنناقش طرق المعادلة الفردية التالية:

1 - المربعات الصغرى العادية (OLS).
2 - المربعات الصغرى غير المباشرة (ILS).
3 - المربعات الصغرى ذات المرحلتين (2SLS).

2.20 النماذج المتزامنة التكرارية والمربعات الصغرى العادية: RECURSIVE MODELS AND ORDINARY LEAST SQUARES

رأينا في الفصل (18)، أنه بسبب التابعية المتبادلة بين مقدار الخطأ العشوائي والمتغير أو المتغيرات المفسرة الداخلية، فإن طريقة OLS لايمكن استخدامها لتقدير معادلة ما في نظام للمعادلات الآنية. فإذا تم تطبيقها عن طريق الخطأ، فإنه كما رأينا في الفقرة 3.18، المقدرات ستكون ليس فقط متحيزة (في العينات الصغيرة)، ولكن أيضًا غير متسقة، بمعنى أن التحيز لايختفي مهما زاد حجم العينة. ولكن عمومًا فإن هناك حالة واحدة يكون من الممكن تطبيق الـ OLS فيها في إطار المعادلات الآنية.

(4) Lawrence R. Klein, A Textbook of Econometrice, 2d ed., Prentice Hall, Englewood Cliffs, N. J., 1974, P. 150.

وهي الحالة الخاصة بالنماذج، المثلثية أو السببية. للتعرف على طبيعة هذه النماذج، دعنا نعتبر النظام ذا المعادلات الثلاث التالية:

$$\begin{aligned} Y_{1t} &= \beta_{10} && + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + u_{1t} \\ Y_{2t} &= \beta_{20} + \beta_{21}Y_{1t} && + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + u_{2t} \\ Y_{3t} &= \beta_{30} + \beta_{31}Y_{1t} + \beta_{32}Y_{2t} && + \gamma_{31}X_{1t} + \gamma_{32}X_{2t} + u_{3t} \end{aligned} \quad (1.2.20)$$

حيث إنه كالعادة، فإن Y's و X's بالترتيب هي المتغيرات الداخلية والخارجية، ومقادير التشتت (الأخطاء) يتحقق فيها التالي:

$$\text{cov}(u_{1t}, u_{2t}) = \text{cov}(u_{1t}, u_{3t}) = \text{cov}(u_{2t}, u_{3t}) = 0$$

بمعنى أن أخطاء نفس الفترة الزمنية في معادلات مختلفة غير مرتبطة (فنيًا فإن هذا هو فرض الارتباط المتعاصر الصغرى).

والآن دعنا نعتبر المعادلة الأولى في (1.2.20). حيث إنها تحتوي فقط على المتغيرات الخارجية على الجانب الأيمن، وبما أن هناك فرض عدم وجود أي ارتباط بينها وبين مقدار الخطأ u_{1t} فإن هذه المعادلة يتحقق فيها الشرط الرئيس لطريقة الـ OLS التقليدية، وهو عدم وجود أي ارتباط بين المتغيرات المفسرة، ومقدار الخطأ العشوائي وبالتالي فإن الـ OLS يمكن تطبيقها مباشرة على هذه المعادلة. اعتبر الآن المعادلة الثانية الموجودة في (1.2.20) والتي تحتوى على المتغير الداخلي Y_1 كمتغير مفسر، بالإضافة إلى المقدار الثابت X's. الآن يمكن تطبيق OLS أيضًا على هذه المعادلة، حيث إن Y_{1t} و u_{2t} غير مرتبطين. هل هذا صحيح؟ الإجابة هي نعم، حيث إن u_1، والتي تؤثر على Y_1، هي مفترض أنها غير مرتبطة مع u_2، وبالتالي لكل الأسباب العملية، فإن Y_1 هو متغير محدد مسبقًا مثله مثل Y_2. وبالتالي يمكن استخدام تقدير الـ OLS لهذه المعادلة. وباستخدام نفس الفكرة من الممكن تطبيق OLS للمعادلة الثالثة في (1.2.20) حيث إن Y_1 و Y_2 غير مرتبطين مع u_3.

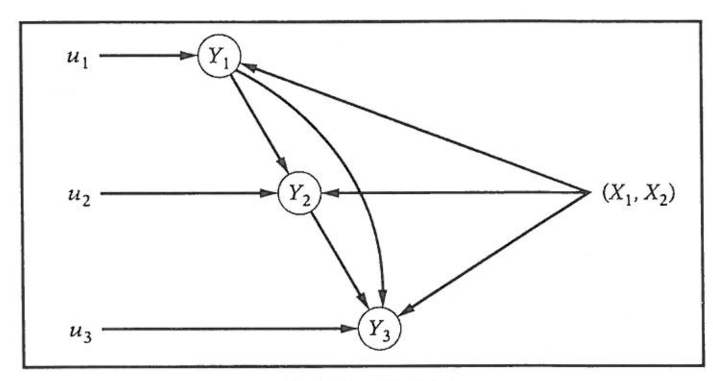
وبالتالي، في التكراري، فإن OLS يمكن تطبيقها لكل معادلة منفردة داخل النظام. فبالفعل لايوجد لدينا مشكلة المعادلات الآنية في مثل هذا الموقف. فمن الشكل البنائي لمثل هذه النماذج، من الواضح أنه لايوجد تابعية متبادلة بين المتغيرات الداخلية. حيث إن Y_1 تؤثر على Y_2، ولكن Y_2 لا تؤثر على Y_1، وبالمثل Y_1 و Y_2 يؤثران على Y_3 بدون أن يتأثرا بها. بمعنى آخر، فإن كل معادلة تمثل تابعية سببية

أحادية الجانب، ومن هنا جاءت تسمية النماذج السببية[5]. تخطيطيًا لدينا شكل (1.20) لتوضيح مثل هذه النماذج.

كمثال للنظام التكراري، دعنا نفترض النموذج التالي للأجر، وتحديد السعر:

معادلة السعر: $\dot{P}_t = \beta_{10} + \beta_{11}\dot{W}_{t-1} + \beta_{12}\dot{R}_t + \beta_{13}\dot{M}_t + \beta_{14}\dot{L}_t + u_{1t}$ **(2.2.20)**

معادلة الأجر: $\dot{W}_t = \beta_{20} + \beta_{21}\text{UN}_t + \beta_{32}\dot{P}_t + u_{2t}$



شكل (1.20) نموذج

حيث إن:

$\dot{P}$ = معدل التغير في السعر لكل وحدة من الناتج.

$\dot{W}$ = معدل التغير في الأجر لكل عامل.

$\dot{R}$ = معدل التغير في سعر رأس المال.

$\dot{L}$ = معدل التغير في إنتاجية العمالة.

UN = معدل البطالة، %[6].

---

(5) الاسم البديل النموذج المثلثي يأتي من فكرة تكوين مصفوفة معاملات المتغيرات الداخلية المعطاة في (1.2.20). فنحصل على المصفوفة المثلثية التالية:

	Y_1	Y_2	Y_3
معادلة 1	1	0	0
معادلة 2	β_{21}	1	0
معادلة 3	β_{31}	β_{32}	1

لاحظ القيم فوق القطر الرئيسي للمصفوفة تساوي الصفر (لماذا؟).

(6) لاحظ أن: النقطة الموجودة فوق الحرف تعني "تفاضل زمني"؟ فمثلاً $\dot{P} = dp/dt$. بالنسبة لسلسلة زمنية متقطعة فإن dp/dt تقرب أحيانًا بـ $\Delta P/\Delta t$، حيث إن الرمز Δ هو المعامل التفاضلي الأول والذي تم مناقشته في الفصل (12).

معادلة السعر تفترض أن معدل التغير في السعر في الفترة الحالية هو دالة في معدل التغير في سعر رأس المال وسعر المواد الخام، معدل التغير في إنتاجية العمالة ومعدل الأجر توضح أن معدل التغير في الأجر في الفترة الحالية يتحدد بمعدل التغير في السعر في الفترة الحالية ومعدل البطالة. من الواضح أن العلاقة السببية تسير في اتجاه $\dot{W}_{t-1} \rightarrow \dot{P}_t \rightarrow \dot{W}_t$ وبالتالي فمن تطبيق الـ OLS لتقدير معلمات المعادلتين فرديًا.

وعلى الرغم من فائدة النماذج التكرارية، فإن غالبية نماذج المعادلات الآنية لايوجد فيها هذا النوع من علاقات السبب، والنتجة أحادية الجانب. وبالتالي فإن OLS في العموم غير مناسبة لتقدير معادلة واحدة منفردة في إطار نموذج المعادلات الآنية[7]. بعض الأفراد يعتقدون أنه على الرغم من أن OLS عمومًا لايمكن تطبيقها في نماذج المعادلات الآنية، فإن الفرد يمكنه استخدامها كبداية وقياس للمقارنة. بمعنى أنه يمكن تقدير المعادلة البنائية باستخدام OLS مع اعتبار خصائص النتائج المتعلقة بالتحيز وعدم الاتساق وغيره من تلك الخصائص. ثم نقدر المعادلة بأي طريقة أخرى صممت خصيصًا للتعامل مع مشكلة الآنية، ويقوم الباحث بمقارنة النتائج التي حصل عليها من كل من الطريقتين على الأقل مقارنة كمية. في العديد من التطبيقات تكون نتائج طريقة OLS غير المناسبة لاتختلف كثيرًا عن مثيلها الذي تم الحصول عليه من الطرق الأكثر تعقيدًا، كما سنرى لاحقًا. فمبدئيًا لابد ألا يعتمد الباحث كثيرًا على النتائج التي سيحصل عليها من طريقة OLS إلا إذا كانت النتائج الحاصل عليها من طرق أخرى متوافرة لدى الباحث. ففي الحقيقة سيعطينا هذا الأسلوب البديل فكرة عن مدى سوء نتائج الـ OLS عندما تطبق في مجال لاتصلح للتطبيق فيه[8].

(7) من المهم أن نضع في الاعتبار وأننا نفترض أن مقدار الخطأ عبر المعادلات المختلفة غير مرتبط متعاصريًا. إذا لم تكن تلك هي الحالة فقد نحتاج إلى استخدام أسلوب التقدير Zellner Sure (انحدارات تبدو غير مرتبطة) لتقدير معالمات النظام المتزامن ذي الاتجاه الواحد. انظر
A. Zellener, "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Assoriation, vol.57. 1962, pp.348-368.

(8) يمكن ملاحظة أنه في حجم العينة الصغير فإن المقدرات البديلة تكون متحيزة مثلها مثل مقدرات الـ OLS ولكن مقدر الـ OLS له خاصية التباين الأقل بين كل المقدرات الأخرى البديلة. ولكن هذا صحيح فقط في العينات الصغيرة.

3.20 تقدير المعادلة تامة التوصيف.. طريقة المربعات الصغرى غير المباشرة : ESTIMATION OF A JUST IDENTIFIED EQUATION- THE METHOD OF INDIRECT LEAST SQUARES (ILS)

الطريقة التي تستخدم في تقدير المعاملات البنائية من تقديرات الـ OLS للمعاملات المخفضة في حالة إذا كانت المعادلة تامة التوصيف تسمى طريقة المربعات الصغرى غير المباشرة (ILS)، والتقديرات التي يتم الحصول عليها تسمى تقديرات المربعات الصغرى غير المباشرة ILS تتم من خلال الخطوات الثلاث التالية :

الخطوة 1 : نحصل أولاً على المعادلات المخفضة الشكل. كما لاحظنا في الفصل (19)، فإن هذه المعادلات المخفضة نحصل عليها من المعادلات البنائية، حيث يكون المتغير التابع في كل معادلة هو متغير داخلي فقط، ودالة في المتغيرات سابقة التحديد (سواء متغيرات خارجية أو داخلية في فترات زمنية متأخرة) والمقدار العشوائي للخطأ فقط.

الخطوة 2 : نطبق الـ OLS على المعادلات المخفضة بشكل منفرد. هذه العملية مسموح بها، حيث إن المتغيرات المفسرة في هذه المعادلات سابقة التحديد مهما جعلها غير مرتبطة مع مقدار الخطأ العشوائي. وبالتالي فالمقدرات التي يتم الحصول عليها تكون مقدرات متسقة[9].

الخطوة 3 : نحصل على مقدرات للمعاملات البنائية الأصلية من مقدرات المعاملات المخفضة من الخطوة 2. وكما لاحظنا في الفصل (19)، إذا كانت المعادلة تامة التوصيف تكون هناك قيمة واحدة مناظرة بين المعاملات البنائية والمعاملات المخفضة، مما يجعلنا نحصل على مقدرات وحيدة للمعاملات الأولى من الأخيرة.

كما رأينا من الخطوات الثلاث السابقة، فقد جاءت تسمية (ILS) من حقيقة أن المعاملات البنائية (الهدف الرئيسي في معظم الحالات) نحصل عليها بشكل غير مباشر من تقديرات الـ OLS للمعاملات المخفضة الشكل.

---

(9) بالإضافة على الاتساق ، فإن المقدرات "قد تكون أفضل المقدرات غير المتحيزة أو مقدرات كافية تقاربيًا أو كلاهما معًا وذلك يعتمد بالترتيب على ما إذا كان *(i)* للمتغير Z= *(X's)* متغير خارجي وليس سابق التحديد (بمعنى ألا يكون متغيراً داخليًا في فترة زمنية متأخرة) و*(ii)* توزيع الخطأ هو التوزيع الطبيعي (W.C. Hood and Tjalling C. Koopmans, Studies in Econometric Method, John Wiley & Sons, New York, 1953, p.133.)

مثال توضيحي : An Illustrative Example

اعتبر نموذج الطلب والعرض المقدم في الفقرة 2.19، وقد تم تبسيط رموزه للتيسير كالتالي :

(1.3.20) دالة الطلب : $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + u_{1t}$

(2.3.20) دالة العرض : $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$

بحيث إن : Q = الكمية

P = السعر

X = الدخل أو النفقات

افترض أن X متغير خارجي. كما سبق وذكرنا، فإن دالة العرض هي دالة تامة التوصيف، في حين أن دالة الطلب غير موصفة. المعادلات المخفضة المرتبطة بالمعادلات البنائية السابقة هي :

$$P_t = \Pi_0 + \Pi_1 X_t + w_t \qquad \textbf{(3.3.20)}$$

$$Q_t = \Pi_2 + \Pi_3 X_t + v_t \qquad \textbf{(4.3.20)}$$

حيث إن Π's هي المعاملات المخفضة الشكل، وهي توليفة غير خطية في المعاملات البنائية كما يتضح من المعادلة (16.2.19) و (18.2.19) حيث إن v, w هما توليفتان خطيتان من مقادير الأخطاء البنائية u_1 و u_2.

لاحظ أن كل معادلة مخفضة الشكل تحتوي على متغير داخلي واحد، وهو متغير تابع، عبارة عن دالة في متغير خارجي X (الدخل) ومقدار خطأ عشوائي فقط.

وبالتالي، فإن معلمات المعادلات مخفضة الشكل السابقة يمكن تقديرها باستخدام الـ OLS وهذه التقديرات هي :

$$\hat{\Pi}_1 = \frac{\sum p_t x_t}{\sum x_t^2} \qquad \textbf{(5.3.20)}$$

$$\hat{\Pi}_0 = \bar{P} - \hat{\Pi}_1 \bar{X} \qquad \textbf{(6.3.20)}$$

$$\hat{\Pi}_3 = \frac{\sum q_t x_t}{\sum x_t^2} \qquad \textbf{(7.3.20)}$$

$$\hat{\Pi}_2 = \bar{Q} - \hat{\Pi}_3 \bar{X} \qquad \textbf{(8.3.20)}$$

حيث إن الحروف الصغيرة كالعادة تعبر عن الانحرافات عن متوسط العينة و $\bar{Q}$ و $\bar{P}$ هي متوسطات محسوبة من العينة لـ Q و P. كما سبق ولاحظنا فإن الـ $\hat{\Pi}_i$'s مقدرات متسقة، وبافتراض بعض الفروض المنطقية تكون أيضًا مقدرات غير متحيزة ذات التباين الأقل أو مقدرات كافية تقاربيًا (انظر الملاحظة 9).

حيث ،إن هدفنا الأساسي هو تحديد المعاملات البنائية، دعنا نرى إذا كنا نستطيع تقديرها من المعاملات المخفضة. كما سبق ورأينا في الفقرة 2.19 فإن دالة العرض هي دالة تامة التوصيف، وبالتالي فمعلمات هذه المعادلة يمكن تقديرها بقيم وحيدة من المعاملات المخفضة كالتالي:

$$\beta_0 = \Pi_2 - \beta_1\Pi_0 \qquad \text{و} \qquad \beta_1 = \frac{\Pi_3}{\Pi_1}$$

وبالتالي، فإن تقديرات هذه المعلمات يمكن الحصول عليها من المعاملات المخفضة كالتالي:

$$\hat{\beta}_0 = \hat{\Pi}_2 - \hat{\beta}_1\hat{\Pi}_0 \qquad (9.3.20)$$

$$\hat{\beta}_1 = \frac{\hat{\Pi}_3}{\hat{\Pi}_1} \qquad (10.3.20)$$

وهذه التقديرات هي تقديرات الـ ILS. لاحظ أن معلمات دالة الطلب لا يمكن تقديرها (انظر تمرين 13.20).

جدول (1.20) إنتاج المحصول ،أسعار المحصول ومصاريف الاستهلاك لكل فرد ،1982 دولاراً ، الولايات المتحدة 1970–1991

Year	Index of crop production (1977 = 100), Q	Index of crop prices received by farmers (1977 = 100), P	Real per capita personal consumption expenditure, X
1970	77	52	3,152
1971	86	56	3,372
1972	87	60	3,658
1973	92	91	4,002
1974	84	117	4,337
1975	93	105	4,745
1976	92	102	5,241
1977	100	100	5,772
1978	102	105	6,384
1979	113	116	7,035
1980	101	125	7,677
1981	117	134	8,375
1982	117	121	8,868
1983	88	128	9,634
1984	111	138	10,408

1985	118	120	11,184
1986	109	107	11,843
1987	108	106	12,568
1988	92	126	13,448
1989	107	134	14,241
1990	114	127	14,996
1991	111	130	15,384

المصدر: Economic report of the president, 1993
بيانات عن Q (جدول B-99)، و P (جدول B-96) و عن X (جدول B-5)

للتعامل مع نتائج رقمية حصلنا على البيانات المعطاة في جدول (1.20). أولاً نقدر المعادلات المخفضة الشكل، ثم نقوم بانحدار السعر والكمية كل منهما منفرداً على مصاريف الاستهلاك الفردية. النتائج كالتالي:

$$\begin{aligned}\hat{P}_t &= 72.3091 + 0.0043X_t \\ se &= (9.2002) \quad (0.0009) \\ t &= (7.8595) \quad (4.4104) \qquad R^2 = 0.4930\end{aligned} \tag{11.3.20}$$

$$\begin{aligned}\hat{Q}_t &= 84.0702 + 0.0020X_t \\ se &= (4.8960) \quad (0.0005) \\ t &= (17.1711) \quad (3.7839) \qquad R^2 = 0.4172\end{aligned} \tag{12.3.20}$$

باستخدام (9.3.20) و (10.3.20) نحصل على مقدرات ILS كالتالي:

$$\hat{\beta}_0 = 51.0562 \tag{13.3.20}$$

$$\hat{\beta}_1 = 0.4566 \tag{14.3.20}$$

وبالتالي، فإن انحدار ILS المقدر هو[10]:

$$\hat{Q}_t = 51.0562 + 0.4566P_t \tag{15.3.20}$$

للمقارنة، دعنا نستعرض نتائج الـ OLS (المطبقة في غير موضعها السليم) لانحدار Q على P:

(10) لم نعط الأخطاء القياسية لمقدرات المعاملات البنائية لأنه، كما لاحظنا من قبل، هذه المعاملات تكون عموماً دوالي غير خطية في المعاملات المخفضة، وبالتالي لاتكون صورة بسيطة لتقدير أخطائهم القياسية من الأخطاء القياسية الخاصة بالمعاملات المخفضة. عندما يكون حجم العينة كبيراً فإنه عموماً يمكن الحصول على قيم تقريبية للأخطاء القياسية للمعاملات البنائية. لمزيد من التفاصيل انظر:
Jan Kementa, Elements of econometrics, Macimillan, New York, 1971, p.444.

$$\hat{Q}_t = 65.1719 + 0.3272 P_t$$
$$\text{se} = (9.3294) \quad (0.0835) \qquad (16.3.20)$$
$$t = (6.9856) \quad (3.9203) \qquad R^2 = 0.4345$$

هذه النتائج توضح كيف أن OLS تعطي نتائج وصورة غير سليمة على الإطلاق، عندما يتم تطبيقها في غير موضعها.

خصائص مقدرات الـ ILS : Properties of ILS Estimators

رأينا أن مقدرات المعاملات المخفضة هي مقدرات متسقة، وبفرض صحة بعض الفروض المنطقية تكون هذه المقدرات أفضل المقدرات غير المتحيزة أو كافية تقاربيًا (انظر الملاحظة 9). هل هذه الخصائص متوافرة أيضًا في مقدرات الـ OLS؟ من الممكن إثبات أن مقدرات ILS يتحقق فيها كل الخصائص التقاربية للمقدرات المخفضة الشكل مثل الاتساق والكفاية التقاربية. لكن (في الأحجام الصغيرة للعينات)، فإن خصائص مثل عدم التحيز ليست بالضرورة متحققة. في الملحق A 20 فقرة 1.A 20 موضحًا أن مقدرات ILS الخاصة بـ $\hat{\beta}_0$ و $\hat{\beta}_1$ لدالة العرض المعطاة سابقًا مقدرات متحيزة، ولكن هذا التحيز يختفي مع زيادة حجم العينة (بمعنى أن هذه المقدرات مقدرات متسقة).(11)

4.20 تقدير المعادلة الموصفة بأكثر مما يجب.. طريقة المربعات الصغرى ذات المرحلتين : Estimation of an overidentified equation- The method of two-stage least squares (2 SLS)

اعتبر النموذج التالي :

(1.4.20) دالة الدخل :
$$Y_{1t} = \beta_{10} + \qquad + \beta_{11} Y_{2t} + \gamma_{11} X_{1t} + \gamma_{12} X_{2t} + u_{1t}$$

(2.4.20) دالة عرض المال :
$$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} \qquad + u_{2t}$$

---

(11) بديهيًا يمكن إثبات ذلك كالتالي : $E(\hat{\beta}_1) = \beta_1$ إذا كان $(\Pi_3 / \Pi_1) = E(\hat{\Pi}_3 / \hat{\Pi}_1)$. والآن حتى إذا كان $E(\hat{\Pi}_3) = \Pi_3$ و $E(\hat{\Pi}_1) = \Pi_3$ فإنه من الممكن إثبات أن $E(\hat{\Pi}_3 / \hat{\Pi}_1) \neq E(\hat{\Pi}_3) / E(\hat{\Pi}_1)$ بمعنى أن توقع خارج قسمة متغيرين لاتساوي خارج قسمة توقع المتغيرين. عمومًا كما هو مثبت في ملحق 1.A20 فإن $\text{plim}\,(\hat{\Pi}_3 / \hat{\Pi}_1) = \text{plim}\,(\hat{\Pi}_3) / \text{plim}(\hat{\Pi}_1) = \Pi_3 / \Pi_1$ بما أن $\hat{\Pi}_3$ و $\hat{\Pi}_1$ مقدرات متسقة.

حيث إن: Y_1 = الدخل

Y_2 = مخزون المال

X_1 = مصاريف الاستثمار

X_2 = مصاريف حكومية على السلع والخدمات

المتغيرات X_1 و X_2 متغيرات خارجية.

معادلة الدخل المولدة من أسلوب Keynesian لنظرية الكمية لتحديد الدخل تنص على أن الدخل يحدد بكل من المعروض من المال، مصاريف الاستثمار والمصاريف الحكومية. دالة عرض المال تفترض أن المخزون من المال محدد (بالنظام الاحتياطي الفيدرالي) على أساس مستوى الدخل. والآن من الواضح أن لدينا مشكلة المعادلات الآنية، ويمكن التأكد من وجودها باستخدام اختبار الآنية الذي تم مناقشته في الفصل (19).

بتطبيق شرط الترتيب للتوصيف، يمكن أن نثبت أن معادلة الدخل موصفة بأقل مما يجب، في حين أن معادلة عرض المال موصفة بأكثر مما يجب. وبالتالي لايوجد شيء نستطيع عمله لتقدير معادلة الدخل إلا إذا تم تغير تعريف النموذج. دالة عرض المال بأكثر مما يجب يمكن ألا يتم تقديرها باستخدام ILS حيث يوجد مقدران لـ β_{21} (يمكن للقارئ إثبات ذلك عن طريق المعاملات المخفضة الشكل).

كنوع من التدريب يمكن تطبيق OLS لمعادلة عرض المال، ولكن المقدرات التي سنحصل عليها ستكون غير متسقة بسبب الارتباط المحتمل بين المتغير العشوائي المفسر Y_1 ومقدار الخطأ العشوائي u_2. افترض أننا وجدنا متغير "مفوض" للمتغير المفسر Y_1 بحيث يكون "مماثلاً" لـ Y_1 (بمعنى أنه مرتبط ارتباطًا قويًا مع Y_1) وغير مرتبط مع u_2. مثل هذا المفوض معروف أيضًا باسم المتغير المساهم (انظر الفصل 17). إذا تم العثور على مثل هذا المفوض فإن OLS يمكن تطبيقها مباشرة لتقدير دالة عرض المال. ولكن كيف يتم الحصول على مثل هذا المتغير المساهم؟ الإجابة عن هذا السؤال تكون من خلال طريقة المربعات الصغرى ذات المرحلتين (2SLS) والتي طورها كل من Heneri Theil[12] و Robert Basman[13] كما يتضح من اسم الطريقة، فإنها تتعلق بتطبيق الـ OLS على مرحلتين، هذه الطريقة تتم كالتالي:

(12)Henri Theil, "Repeated Least-Squares Applied to Complete Equation Systems," The Hague: The Central Planning Bureau, The Netherlands, 1953 (mimeographed).

(13) Robert L. Basmann, "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," Econometrica, vol. 25, 1957, pp. 77-83.

المرحلة 1: للتخلص من الارتباط المحتمل بين Y_1 و u_2، فإننا نقوم أولاً بعمل انحدار لـ Y_1 على كل المتغيرات المحددة سابقًا في النظام كله، وليس فقط المتغيرات الموجودة في المعادلة. في المثال الحالي، فإن هذا يعني انحدار Y_1 على X_1 و X_2 معًا كالتالي:

$$Y_{1t} = \hat{\Pi}_0 + \hat{\Pi}_1 X_{1t} + \hat{\Pi}_2 X_{2t} + \hat{u}_t \qquad \textbf{(3.4.20)}$$

بحيث إن $\hat{u}_t$ هو تقدير OLS التقليدي. من المعادلة (3.4.20) نحصل على:

$$\hat{Y}_{1t} = \hat{\Pi}_0 + \hat{\Pi}_1 X_{1t} + \hat{\Pi}_2 X_{2t} \qquad \textbf{(4.4.20)}$$

حيث إن $\hat{Y}_{1t}$ هو تقدير للقيمة المتوسطة لقيمة Y المشروطة على قيم ثابتة لـ X's، لاحظ أن (3.4.20) ليست إلا انحدارًا مخفض الشكل، حيث لايظهر في الجانب الأيمن للمعادلة إلا متغيرات خارجية أو سابقة التحديد. معادلة (4.3.20) يمكن التعبير عنها كالتالي:

$$Y_{1t} = \hat{Y}_{1t} + \hat{u}_t \qquad \textbf{(5.4.20)}$$

والتي توضح أن المتغير العشوائي Y_1 يشمل جزءين: $\hat{Y}_{1t}$ والذي يعتبر توليفة خطية من المتغير غير العشوائي X's وجزء عشوائي $\hat{u}_t$.

وفقًا لنظرية OLS، فإن $\hat{Y}_{1t}$ و $\hat{u}_t$ غير مرتبطين (لماذا؟).

المرحلة 2: معادلة عرض المال الموصفة بأكثر مما يجب، يمكن كتابتها كالتالي:

$$\begin{aligned} Y_{2t} &= \beta_{20} + \beta_{21}(\hat{Y}_{1t} + \hat{u}_t) + u_{2t} \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + (u_{2t} + \beta_{21}\hat{u}_t) \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + u_t^* \end{aligned} \qquad \textbf{(6.4.20)}$$

حيث إن:

$$u_t^* = u_{2t} + \beta_{21}\hat{u}_t$$

بمقارنة (6.4.20) مع (2.4.20) نرى أنهما متقاربان في الشكل، والفرق الوحيد هو أن Y_1 مستبدلة بـ $\hat{Y}_1$. فما هي فائدة (6.4.20)؟ يمكن إثبات أنه على الرغم من أن Y_1 في معادلة عرض المال الأصلية مرتبطة أو محتمل ارتباطها مع مقدار الخطأ u_2 (وبالتالي تطبيق OLS يكون غير مناسب) فإن $\hat{Y}_{1t}$ في (6.4.20) غير مرتبطة تقاربيًا مع

u_t^*، أي أن ذلك يتحقق مع حجم العينة الكبير (أو بشكل أدق مع زيادة حجم العينة)، مما سيعطي مقدرات متسقة لمعلمات دالة عرض المال[14].

كما يتضح من هذه الطريقة ذات المرحلتين، فإن الفكرة الرئيسة وراء 2SLS هي "تقنية" المتغير العشوائي المفسر Y_1 من أثر مقدار الخطأ العشوائي u_2. هذا الهدف يتحقق بتطبيق انحدار مخفض الشكل لـ Y_1 على كل المتغيرات المحددة سابقًا في النظام (المرحلة 1) فنحصل على تقديرات لـ$\hat{Y}_{1t}$ ونصفها بدلاً من Y_{1t} في المعادلة الأصلية ثم نطبق OLS على هذه المعادلة التي تم تحويلها (مرحلة 2). المقدرات التي سيتم الحصول عليها ستكون متسقة، بمعنى أنها تؤول إلى قيمها الحقيقية مع زيادة حجم العينة.

لشرح طريقة 2SLS بشكل أفضل، دعنا نعدل نموذج عرض المال- الدخل كالتالي:

$$Y_{1t} = \beta_{10} + \beta_{12}Y_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + u_{1t} \quad (7.4.20)$$

$$Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} + \gamma_{23}X_{3t} + \gamma_{24}X_{4t} + u_{2t} \quad (8.4.20)$$

فبالإضافة للمتغيرات السابق تعريفها، فإن X_3 = الدخل في الفترة الزمنية السابقة، و X_4 = المعروض من المال في الفترة الزمنية السابقة. كل من X_3 و X_4 سابقة التحديد. من الممكن إثبات أن كلاً من المعادلتين (7.4.20) و (8.4.20) هما معادلتان موصفتان بأكثر مما يجب. لتطبيق 2SLS نقوم بالتالي: في المرحلة 1 نقوم بانحدار المتغيرات الداخلية على كل المتغيرات سابقة التحديد في النظام كالتالي:

$$Y_{1t} = \hat{\Pi}_{10} + \hat{\Pi}_{11}X_{1t} + \hat{\Pi}_{12}X_{2t} + \hat{\Pi}_{13}X_{3t} + \hat{\Pi}_{14}X_{4t} + \hat{u}_{1t} \quad (9.4.20)$$

$$Y_{2t} = \hat{\Pi}_{20} + \hat{\Pi}_{21}X_{1t} + \hat{\Pi}_{22}X_{2t} + \hat{\Pi}_{23}X_{3t} + \hat{\Pi}_{24}X_{4t} + \hat{u}_{2t} \quad (10.4.20)$$

في المرحلة 2 نستبدل Y_1، Y_2 في المعادلات الأصلية (البنائية) بقيمهما المقدرة من الانحدارين السابقين ثم نقوم بعمل انحدارات OLS كالتالي:

(14) لاحظ أنه في العينات الصغيرة فإن $\hat{Y}_{1t}$ محتمل أن تكون مرتبطة مع u_t^*. السبب في ذلك كالتالي: من المعادلة (4.4.20) نرى أن $\hat{Y}_{1t}$ هي توليفة خطية مرجحة من المتغيرات المحددة سابقًا X's و $\hat{\Pi}$'s هي الأوزان. الآن حتى إذا كانت هذه المتغيرات المحدودة سابقًا غير عشوائية فإن $\hat{\Pi}$'s، كمقدرات تكون عشوائية وبالتالي فإن $\hat{Y}_{1t}$ عشوائية أيضًا. والآن من خلال مناقشتنا للمعادلات المخفضة الشكل وتقدير المربعات الصغرى غير المباشرة من الواضح أن المعاملات المخفضة، $\hat{\Pi}$'s ، هي دوال في مقادير الأخطاء العشوائية مثل u_2. وبما أن $\hat{Y}_{1t}$ تعتمد على $\hat{\Pi}$'s فإنه من المحتمل ارتباطها مع u_t والذي هو مكون من مكونات u_t^*. وكنتيجة لذلك $\hat{Y}_{1t}$ من المتوقع أن تكون مرتبطة مع u_t^* . ولكن كما سبق وذكرنا هذا الارتباط يختفي مع زيادة حجم العينة. خلاصة ذلك أنه عندما يكون حجم العينة صغيراً فإن طريقة 2SLS قد تؤدي إلى مقدرات متحيزة.

$$Y_{1t} = \beta_{10} + \beta_{12}\hat{Y}_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + u^*_{1t} \qquad (11.4.20)$$

$$Y_{2t} = \beta_{20} + \beta_{21}\hat{Y}_{1t} + \gamma_{23}X_{3t} + \gamma_{24}X_{4t} + u^*_{2t} \qquad (12.4.20)$$

حيث إن $u^*_{1t} = u_{1t} + \beta_{12}\hat{u}_{2t}$ و $\hat{u}^*_{2t} = u_{2t} + \beta_{12}\hat{u}_{1t}$. المقدرات التي سيتم الحصول عليها ستكون مقدرات متسقة..

لاحظ الخصائص التالية في طريقة 2SLS.

1 - يمكن تطبيق هذه الطريقة على كل معادلة منفردة في النظام بدون أخذ في الاعتبار باقي المعادلات في النظام. وبالتالي لحل نماذج الاقتصاد القياسي التي تشتمل على عدد كبير من المعادلات، فإن 2SLS تعتبر طريقة اقتصادية مناسبة. ولهذا السبب، فإن هذه الطريقة تم استخدامها كثيراً في التطبيقات العملية.

2 - على عكس ILS والتي تعطي تقديرات متعددة للمعلمات في حالة المعادلات الموصفة بأكثر مما يجب، فإن 2SLS تعطي تقديرات وحيدة لكل معلمة.

3 - من السهل تطبيق هذه الطريقة، حيث إن كل مانحتاج لمعرفته هو العدد الإجمالي للمتغيرات الخارجية أو السابقة تحديد في النظام بدون الحاجة لمعرفة أي متغيرات أخرى في النظام.

4 - على الرغم من أن هذه الطريقة صممت خصيصاً للتعامل مع المعادلات الموصفة بأكثر مما يجب، فإن هذه الطريقة يمكن تطبيقها أيضاً للمعادلات تامة التوصيف ولكن هنا ستكون مقدرات ILS هي نفسها مقدرات الـ 2SLS. (لماذا؟)

5 - إذا كانت قيم R^2 في الانحدارات المخفضة الشكل (أي انحدارات المرحلة 1) ذات قيم كبيرة، مثلاً، أكبر من 0.8 فإن مقدرات OLS التقليدية ومقدرات 2SLS ستكون متقاربة جداً. هذه النتيجة تكون متوقعة إذا كانت قيمة R^2 في المرحلة الأولى عالية جداً، حيث إن ذلك يعني أن القيم المقدرة للمتغيرات الداخلية قريبة جداً من قيمها الحقيقية، وبالتالي فالأخيرة يكون احتمالها أقل للارتباط مع الأخطاء العشوائية في المعادلات البنائية الأصلية. (لماذا)[15].

(15) في الحالة القصوى إذا كانت $R^2=1$ في الانحدار في المرحلة الأولى، فإن المتغير المفسر الداخلي في المعادلة الأصلية (الموصفة بأكثر مما يجب) يكون عملياً غير عشوائي (لماذا؟).

عمومًا إذا كانت قيمة R^2 في انحدار المرحلة الأولى منخفضة جداً، فإن مقدرات 2SLS ستكون عمليًا بدون معنى، حيث إننا عند تبديل الـ Y's الأصلية في انحدار المرحلة الثانية بالقيمة المقدرة لـ $\hat{Y}$ s من انحدار المرحلة الأولى، سيكون هذا الأخير ممثلاً للخطأ في انحدار المرحلة الأولى. بمعنى آخر، في مثل هذه الحالة، فإن $\hat{Y}$ s ستكون مفوضًا ضعيفًا جداً عن Ys الأصلية.

6 - لاحظ أنه عند كتابة تقرير نتائج انحدار ILS في (15.3.20) لم نذكر الأخطاء القياسية للمعادلات المقدرة (للأسباب المشروحة في الملاحظة 10). ولكننا نستطيع عمل ذلك بالنسبة لمقدرات 2SLS، حيث إن المعاملات البنائية مقدرة مباشرة من انحدار OLS في المرحلة الثانية. وعمومًا هناك شئ لابد من أخذه في الاعتبار، فالأخطاء القياسية المقدرة في انحدار المرحلة الثانية، تحتاج إلى بعض التعديل، حيث كما نرى في المعادلة (6.4.20) فمقدار الخطأ u_t^* ماهو في الحقيقة إلا مقدار الخطأ الأصلي u_{2t} بالإضافة إلى $\beta_{21}\hat{u}_t$، وبالتالي فإن تباين u_t^* ليس بالضبط مساويًا لتباين u_{2t} الأصلي. عمومًا هذا التعديل يمكن فهمه بسهولة من المعادلة المعطاة في الملحق A20، فقرة 20 2.A.

7 - باستخدام الـ 2SLS، ضع في الاعتبار الملاحظات التالية لـ Henri Theil التعليل الإحصائي لطريقة 2SLS في العينات ذات الأحجام الكبيرة. عندما لا توجد متغيرات داخلية في فترات زمنية متأخرة، . . فالمعاملات المقدرة عن طريق الـ 2SLS تكون متسقة، إذا كانت المتغيرات الخارجية ثابتة في العينات المكررة، وإذا كان مقدار الخطأ (الذي يظهر في المعادلات البنائية) مستقلاً وموزعًا منفردًا بتوقع يساوي الصفر وتباين محدود. إذا تحقق هذان الشرطان فإن توزيع المعاملات المقدرة باستخدام 2SLS يكون له التوزيع الطبيعي تقاربيًا للعينات الكبيرة.

عندما يحتوي نظام المعادلات على متغيرات داخلية في فترات زمنية متأخرة، فإن الاتساق واتباع التوزيع الطبيعي في العينات كبيرة الحجم لمقدرات المعاملات بطريقة 2SLS تحتاج إلى شرط آخر إضافي. وهو بزيادة حجم العينة، فإن متوسط مربعات القيم المأخوذة للمتغير الداخلي في فترات زمنية متأخرة تقترب احتماليًا لنهاية ما متوقعة.

إذا كان (مقدار الخطأ الظاهر في المعادلات البنائية المختلفة) غير مستقل، فإن قيم المتغيرات الداخلية في فترات زمنية متأخرة تكون غير مستقلة عن العمليات الحالية

لنظام المعادلات، مما يعني أن هذه المتغيرات ليست بالفعل سابقة التحديد. وإذا لم تتم معاملة هذه المتغيرات على أنها متغيرات سابقة التحديد، فإنه في طريقة الـ 2SLS فإن المقدرات المستنتجة تكون مقدرات غير متسقة[16].

5.20 2SLS: مثال رقمي : 2SLS: A NUMERICAL EXAMPLE

لشرح طريقة الـ 2SLS، دعنا نعتبر نموذج عرض المال- الدخل المعطى سابقًا في المعادلتين (1.4.20) و (2.4.20)، كما هو موضح سابقًا، فإن معادلة عرض المال هي معادلة موصفة بأكثر مما يجب. لتقدير معلمات هذه المعادلة، فإننا نلجأ إلى طريقة المربعات الصغرى ذات المرحلتين. البيانات المطلوبة لهذا التحليل معطاة في جدول (2.20)، هذا الجدول يعطي أيضًا البيانات المطلوب للإجابة على بعض الأسئلة الموجودة في التمارين.

انحدار المرحلة 1: أولاً انحدار للمتغير العشوائي المفسر (الدخل) Y_1 والممثل بـ GDP على المتغير المحدد سابقًا (الاستثمار الخاص) X_1 و المصاريف الحكومية X_2، فنحصل على النتائج التالية:

جدول (2.20) GDP، M2، GPDI، FEDEXP، TB6، الولايات المتحدة 1970–1999

YEAR	GDP (Y_1)	M2 (Y_2)	GPDI (X_1)	FEDEXP (X_2)	TB6 (X_3)
1970	3578.000	626.4000	436.2000	198.6000	6.562000
1971	3697.700	710.1000	485.8000	216.6000	4.511000
1972	3998.400	802.1000	543.0000	240.0000	4.466000
1973	4123.400	855.2000	606.5000	259.7000	7.178000
1974	4099.000	901.9000	561.7000	291.2000	7.926000
1975	4084.400	1015.900	462.2000	345.4000	6.122000
1976	4311.700	1151.700	555.5000	371.9000	5.266000
1977	4511.800	1269.900	639.4000	405.0000	5.510000
1978	4760.600	1365.500	713.0000	444.2000	7.572000
1979	4912.100	1473.100	735.4000	489.6000	10.01700
1980	4900.900	1599.100	655.3000	576.6000	11.37400
1981	5021.000	1754.600	715.6000	659.3000	13.77600
1982	4913.300	1909.500	615.2000	732.1000	11.08400
1983	5132.300	2126.000	673.7000	797.8000	8.750000
1984	5505.200	2309.700	871.5000	856.1000	9.800000
1985	5717.100	2495.400	863.4000	924.6000	7.660000
1986	5912.400	2732.100	857.7000	978.5000	6.030000
1987	6113.300	2831.100	879.3000	1018.400	6.050000
1988	6368.400	2994.300	902.8000	1066.200	6.920000
1989	6591.900	3158.400	936.5000	1140.300	8.040000
1990	6707.900	3277.600	907.3000	1228.700	7.470000

(16) Henri Theil, Introduction to econometrics, Prentice Hall, Englewood Cliffs, N.J., 1987, pp. 341-243.

1991	6676.400	3376.800	829.5000	1287.600	5.490000
1992	6880.000	3430.700	899.8000	1418.900	3.570000
1993	7062.600	3484.400	977.9000	1471.500	3.140000
1994	7347.700	3499.000	1107.000	1506.000	4.660000
1995	7543.800	3641.900	1140.600	1575.700	5.590000
1996	7813.200	3813.300	1242.700	1635.900	5.090000
1997	8159.500	4028.900	1393.300	1678.800	5.180000
1998	8515.700	4380.600	1566.800	1705.000	4.850000
1999	8875.800	4643.700	1669.700	1750.200	4.760000

لاحظ أن: Y_1 = GDP= الإنتاج الكلي المحلي (بليون دولار، معدلة موسميًا).

Y_2 = M2 = $\overline{M}$ 2 عرض المال (بليون دولار معدلة موسميًا).

X_1 = GDP1 = الاستثمار الكلي المحلي الخاص (بليون دولار، معدلة موسميًا).

X_2 = FEDEXP= مصاريف فيدرالية حكومية (بليون دولار، معدلة موسميًا).

X_3 = TB6= معدل الأوراق النقدية لوزارة المالية كل 6 شهور، (%)

المصدر: جداول B-2، B-96، B-73، Economic report of the president, 2001, 84-B.

$$\hat{Y}_{1t} = 2587.351 + 1.6707X_{1t} + 1.9693X_{2t}$$
$$\text{se} = (72.0011) \quad (0.1646) \quad (0.0983) \qquad \textbf{(1.5.20)}$$
$$t = (35.9349) \quad (10.1489) \quad (20.0200) \qquad R^2 = 0.9947$$

انحدار المرحلة 2: الآن نقدر دالة عرض المال (2.4.20) ونستبدل المتغير الداخلي Y_1 بالقيمة المقدرة Y_1 من (1.5.20) ($=\hat{Y}_1$). النتائج كالتالي:

$$\hat{Y}_{2t} = -2198.297 + 0.7916\hat{Y}_{1t}$$
$$\text{se} = (139.0986) \quad (0.0232) \qquad \textbf{(2.5.20)}$$
$$t = (-15.8038) \quad (34.0502) \qquad R^2 = 0.9764$$

كما أشرنا من قبل، فإن الأخطاء القياسية المقدرة المعطاة في (2.5.20) لابد من تصحيحها، كما هو موضح في الملحق 20 A.2، فقرة 20 A.2. بعد القيام بهذا التعديل (معظم حزم البرامج الإلكترونية الخاصة بالاقتصاد القياسي تقوم بحساب مثل هذا التعديل) نحصل على النتائج التالية:

$$\hat{Y}_{2t} = -2198.297 + 0.7915\hat{Y}_{1t}$$
$$\text{se} = (126.9598) \quad (0.0212) \qquad \textbf{(3.5.20)}$$
$$t = (-17.3149) \quad (37.3057) \qquad R^2 = 0.9803$$

كما هو موضح في الملحق A20، فقرة 20 A.2، الأخطاء القياسية المعطاة في (3.5.20) لاتختلف كثيراً عن تلك المعطاة في (2.5.20) وذلك بسبب أن R^2 في المرحلة الأولى تعتبر قيمة عالية جداً.

انحدار OLS : للمقارنة، دعنا نستعرض انحدار المخزون من المال على الدخل، كما هو موضح في (2.4.20) بدون تنقية Y_{1t} العشوائي من أثر مقدار الخطأ العشوائي.

$$\hat{Y}_{2t} = -2195.468 + 0.7911Y_{1t}$$
$$\text{se} = (126.6460) \quad (0.0211) \qquad (4.5.20)$$
$$t = (-17.3354) \quad (37.3812) \qquad R^2 = 0.9803$$

بمقارنة نتائج OLS (غير الصالحة للتطبيق في مثل هذه الحالة) مع انحدار المرحلة 2، نرى أن الانحدارين متقاربان جداً. هل يعني ذلك أن طريقة 2SLS عديمة القيمة؟ لا. في المثال الحالي، فإن النتيجتين متساويتان، وهذا من المتوقع حدوثه، حيث سبق وذكرنا أن قيمة R^2 التي حصلنا عليها في المرحلة 1 قيمة عالية جداً، مما يجعل تقدير $\hat{Y}_{1t}$ مماثلاً للقيمة الحقيقية لـ Y_{1t}. وبالتالي في مثل هذه الحالة يكون انحدار OLS، والانحدار على مرحلتين متقاربتين. ولكن لايوجد ما يضمن ما سيحدث في كل التطبيقات. وبالتالي فتطبيقيًا في حالة المعادلات الموصفة بأكثر مما يجب، يمكن للفرد أن يقبل نتائج OLS التقليدية بدون اللجوء إلى الانحدار ذي المرحلتين.

الآنية بين GDP و عرض المال. دعنا ندرس الآن ما إذا كانت هناك تبعية تبادلية بين GDP (Y_1) والمعروض من المال (Y_2). للقيام بذلك نستخدم اختبار Hausman للآنية والذي سبق مناقشته في الفصل (19)، أولاً نقوم بعمل انحدار لـ GDP على X_1 (مصاريف الاستثمار)، و X_2 (المصاريف الحكومية)، حيث يمثلان المتغيرات الخارجية في النظام (بمعنى أننا نقدر انحدار مخفض الشكل). من هذا الانحدار نحصل على القيم المقدرة لـ GDP والبواقي $\hat{v}_t$. كما هو موجود في المعادلة (7.4.19)، ثم نقوم بعمل انحدار للمعروض من المال على القيمة المقدرة لـ GDP و v_t فنحصل على النتائج التالية:

$$\hat{Y}_{2t} = -2198.297 + 0.7915\hat{Y}_{1t} + 0.6984\hat{v}_t$$
$$\text{se} = (129.0548) \quad (0.0215) \quad (0.2970) \qquad (5.5.20)$$
$$t = (-17.0338) \quad (36.70016) \quad (2.3511)$$

بما أن قيمة t الخاصة بـ $\hat{v}_t$ معنوية إحصائيًا (قيمة P هي 0.0263)، فنحن لانستطيع رفض الفرض الخاص بوجود علاقة آنية بين المعروض من المال و GDP. وهذه النتيجة غير مستبعدة. (لاحظ أن هذا الاستنتاج مقصور على أحجام العينات الكبيرة، فنيًا هذا يصلح فقط مع زيادة حجم العينة).

اختبار الفروض: افترض أننا نريد اختبار الفرض القائل أن الدخل ليس له أي تأثير على طلب المال. هل ممكن القيام بمثل هذا الاختبار باستخدام اختبار t العادي من الانحدار المقدر من (2.5.20)؟ نعم. فبافتراض كبر حجم العينة وبعد القيام بتصحيح الأخطاء القياسية كما موضح في (3.5.20)، يمكن استخدام اختبار t لاختبار معنوية أي معامل بشكل منفرد، واختبار F لاختبار معنوية معاملان أو أكثر معًا، ويتم ذلك باستخدام الصيغة (7.5.8)(17).

ماذا سيحدث إذا كان مقدار الخطأ في المعادلة البنائية مرتبطًا ذاتيًا أو مرتبطًا مع مقدار الخطأ في معادلة بنائية أخرى في النظام؟ الإجابة الكاملة عن هذا السؤال تقع نوعًا ما خارج نطاق هذا الكتاب، ويفضل قراءة المزيد عنها في المراجع (انظر المرجع المعطى في الملاحظة 7). وعمومًا فإن هناك بعض طرق التقدير (مثل طريقة Zellner's Sure) الموجودة للتعامل مع مثل هذا التعقيد.

6.20 أمثلة توضيحية : ILLUSTRATIVE EXAMPLE

في هذه الفقرة، دعنا نستعرض بعض التطبيقات لطرق المعادلات الآنية.

مثال 1.20

الدعاية، التركيز وحدود السعر : Advertising, Concentration, and price margins

لدراسة العلاقات بين الدعاية، التركيز (مقاس بنسبة التركيز) وحدود التكلفة– السعر، قام Auyn D.Strickland و Lenard W.Weiss بتصميم النموذج ثلاثي المعادلات التالي(18).

دالة كثافة الدعاية :

$$Ad/S = a_0 + a_1 M + a_2(CD/S) + a_3 C + a_4 C^2 + a_5 Gr + a_6 Dur \qquad (1.6.20)$$

دالة التركيز :

$$C = b_0 + b_1(Ad/S) + b_2(MES/S) \qquad (2.6.20)$$

(17) لكن ضع في الاعتبار التالي: RSS المقيد وغير المقيد الموجود في البسط يجب حسابه باستخدام قيم Y المتنبأ بها (كما في المرحلة 2 من 2 SLS) والـ RSS الموجود في المقام محسوبة باستخدام القيم الحقيقية بدلاً من المقدرة للمتغيرات المنحدرة. لمناقشة أكثر توسعًا في هذه النقطة، انظر: T.Dudley Wallace and j. lew Silver, Econometrics: An Introduction, Addison-Wesley, Reading, Mass., 1988, sec. 8.5.

(18) انظر: "Advertising, Concentration, and price- cost margins, "Journal of Political Economy", vol. 84, no.5, 1976. pp.1109-1121.

دالة حدود التكلفة- السعر :

$$M = c_0 + c_1(K/S) + c_2\text{Gr} + c_3C + c_4\text{GD} + c_5(\text{Ad}/S) + c_6(\text{MES}/S) \qquad (3.6.20)$$

بحيث إن: Ad= تكلفة الدعاية
S = قيمة الشحنة
C = نسبة التركيز أربع مؤسسات تجارية
CD = طلب المستهلك
MES= الحد الأدنى لمقياس الكفاية
M = حد السعر/ التكلفة
Gr = المعدل السنوي لنمو الإنتاج الصناعي
DUR= متغير وهمي للسلع الصناعية الاستهلاكية
K = مخزون رأس المال
GD = مقياس للإنتشار الجغرافي للناتج

باستخدام شرط الترتيب للتوصيف، فإن المعادلة (2.6.20) موصفة بأكثر مما يجب، في حين (1.6.20) و (3.6.20) تامة التوصيف.

بيانات هذا التحليل تأتي بشكل أساسي من مراكز التعداد 1963، وتغطي 408 مصانع لسلع صناعية ذات أربعة أرقام من 417 مصنعاً موجوداً. المعادلات الثلاث تم تقديرها أولاً باستخدام OLS فحصلنا على النتائج الموجودة في جدول (3.20). لتصحيح التحيز الناشئ من وجود معادلات آنية. قام الباحثون بإعادة تقدير النموذج باستخدام 2SLS، وهذه النتائج موجودة في جدول (4.20). سنترك للقارئ المقارنة بين النتيجتين.

جدول (3.20) تقديرات OLS للمعادلات الثلاث (نسبة t موجودة بين الأقواس)

	Dependent variable		
	Ad/*S* Eq. (20.6.1)	*C* Eq. (20.6.2)	*M* Eq. (20.6.3)
Constant	−0.0314 (−7.45)	0.2638 (25.93)	0.1682 (17.15)
C	0.0554 (3.56)	—	0.0629 (2.89)
C^2	−0.0568 (−3.38)	—	—
M	0.1123 (9.84)	—	—
CD/*S*	0.0257 (8.94)	—	—
Gr	0.0387 (1.64)		0.2255 (2.61)
Dur	−0.0021 (−1.11)	—	—
Ad/*S*	—	1.1613 (3.3)	1.6536 (11.00)
MES/*S*	—	4.1852 (18.99)	0.0686 (0.54)
K/S	—	—	0.1123 (8.03)
GD	—	—	−0.0003 (−2.90)
R^2	0.374	0.485	0.402
df	401	405	401

جدول (4.20) تقديرات المربعات الصغرى ذات المرحلتين للمعادلات الثلاث (نسبة t موجودة بين الأقواس)

	Dependent variable		
	Ad/*S* Eq. (20.6.1)	*C* Eq. (20.6.2)	*M* Eq. (20.6.3)
Constant	−0.0245 (−3.86)	0.2591 (21.30)	0.1736 (14.66)
C	0.0737 (2.84)	—	0.0377 (0.93)
C^2	−0.0643 (−2.64)	—	—
M	0.0544 (2.01)	—	—
CD/*S*	0.0269 (8.96)	—	—
Gr	0.0539 (2.09)	—	0.2336 (2.61)
Dur	−0.0018 (−0.93)	—	—
Ad/*S*	—	1.5347 (2.42)	1.6256 (5.52)
MES/*S*	—	4.169 (18.84)	0.1720 (0.92)
K/S	—	—	0.1165 (7.30)
GD	—	—	−0.0003 (−2.79)

مثال 2.20 :

نموذج Klein's

في مثال 6.18 ناقشنا باختصار نموذج الريادة لـ Klein . مبدئيًا هذا النموذج تم تقديره في الفترة 1920−1941 . البيانات الخاصة بذلك معطاة في جدول (5.20)، وتقديرات OLS مخفضة الشكل وتقديرات 2SLS معطاة في جدول (6.20). متروك للقارئ تفسير هذه النتائج .

جدول (5.20) البيانات الخاصة بنموذج Klein's

Year	C^*	*P*	*W*	*I*	K_{-1}	*X*	*W'*	*G*	*T*
1920	39.8	12.7	28.8	2.7	180.1	44.9	2.2	2.4	3.4
1921	41.9	12.4	25.5	−0.2	182.8	45.6	2.7	3.9	7.7
1922	45.0	16.9	29.3	1.9	182.6	50.1	2.9	3.2	3.9
1923	49.2	18.4	34.1	5.2	184.5	57.2	2.9	2.8	4.7
1924	50.6	19.4	33.9	3.0	189.7	57.1	3.1	3.5	3.8
1925	52.6	20.1	35.4	5.1	192.7	61.0	3.2	3.3	5.5
1926	55.1	19.6	37.4	5.6	197.8	64.0	3.3	3.3	7.0
1927	56.2	19.8	37.9	4.2	203.4	64.4	3.6	4.0	6.7
1928	57.3	21.1	39.2	3.0	207.6	64.5	3.7	4.2	4.2
1929	57.8	21.7	41.3	5.1	210.6	67.0	4.0	4.1	4.0
1930	55.0	15.6	37.9	1.0	215.7	61.2	4.2	5.2	7.7
1931	50.9	11.4	34.5	−3.4	216.7	53.4	4.8	5.9	7.5
1932	45.6	7.0	29.0	−6.2	213.3	44.3	5.3	4.9	8.3
1933	46.5	11.2	28.5	−5.1	207.1	45.1	5.6	3.7	5.4
1934	48.7	12.3	30.6	−3.0	202.0	49.7	6.0	4.0	6.8
1935	51.3	14.0	33.2	−1.3	199.0	54.4	6.1	4.4	7.2
1936	57.7	17.6	36.8	2.1	197.7	62.7	7.4	2.9	8.3
1937	58.7	17.3	41.0	2.0	199.8	65.0	6.7	4.3	6.7
1938	57.5	15.3	38.2	−1.9	201.8	60.9	7.7	5.3	7.4
1939	61.6	19.0	41.6	1.3	199.9	69.5	7.8	6.6	8.9
1940	65.0	21.1	45.0	3.3	201.2	75.7	8.0	7.4	9.6
1941	69.7	23.5	53.3	4.9	204.5	88.4	8.5	13.8	11.6

(*) تفسير عناوين الأعمدة موجود في مثال 6.18

المصدر : هذه البيانات تم الحصول عليها من

G.S. Maddala, Econometrics, Mc Craw- Hill, New York, 1977, p.238.

جدول (6.20)(*) تقديرات OLS مخفضة الشكل وتقديرات 2SLS لنموذج Klein's

OLS:

$$\hat{C} = \underset{(1.203)}{16.237} + \underset{(0.091)}{0.193}P + \underset{(0.040)}{0.796}(W+W') + \underset{(0.090)}{0.089}P_{-1} \qquad \bar{R}^2 = 0.978 \qquad DW = 1.367$$

$$\hat{I} = \underset{(5.465)}{10.125} + \underset{(0.097)}{0.479}P + \underset{(0.100)}{0.333}P_{-1} - \underset{(0.026)}{0.112}K_{-1} \qquad \bar{R}^2 = 0.919 \qquad DW = 1.810$$

$$\hat{W} = \underset{(1.151)}{0.064} + \underset{(0.032)}{0.439}X + \underset{(0.037)}{0.146}X_{-1} + \underset{(0.031)}{0.130}t \qquad \bar{R}^2 = 0.985 \qquad DW = 1.958$$

Reduced-form:

$$\hat{P} = \underset{(10.870)}{46.383} + \underset{(0.444)}{0.813}P_{-1} - \underset{(0.067)}{0.213}K_{-1} + \underset{(0.252)}{0.015}X_{-1} + \underset{(0.154)}{0.297}t - \underset{(0.385)}{0.926}T + \underset{(0.373)}{0.443}G$$
$$\bar{R}^2 = 0.753 \qquad DW = 1.854$$

$$\widehat{W+W'} = \underset{(8.787)}{40.278} + \underset{(0.359)}{0.823}P_{-1} - \underset{(0.054)}{0.144}K_{-1} + \underset{(0.204)}{0.115}X_{-1} + \underset{(0.124)}{0.881}t - \underset{(0.311)}{0.567}T + \underset{(0.302)}{0.859}G$$
$$\bar{R}^2 = 0.949 \qquad DW = 2.395$$

$$\hat{X} = \underset{(18.860)}{78.281} + \underset{(0.771)}{1.724}P_{-1} - \underset{(0.110)}{0.319}K_{-1} + \underset{(0.438)}{0.094}X_{-1} + \underset{(0.267)}{0.878}t - \underset{(0.669)}{0.565}T + \underset{(0.648)}{1.317}G$$
$$\bar{R}^2 = 0.882 \qquad DW = 2.049$$

2SLS:

$$\hat{C} = \underset{(1.464)}{16.543} + \underset{(0.130)}{0.019}P + \underset{(0.044)}{0.810}(W+W') + \underset{(0.118)}{0.214}P_{-1} \qquad \bar{R}^2 = 0.9726$$

$$\hat{I} = \underset{(8.361)}{20.284} + \underset{(0.191)}{0.149}P + \underset{(0.180)}{0.616}P_{-1} - \underset{(0.040)}{0.157}K_{-1} \qquad \bar{R}^2 = 0.8643$$

$$\hat{W} = \underset{(1.894)}{0.065} + \underset{(0.065)}{0.438}X + \underset{(0.070)}{0.146}X_{-1} + \underset{(0.053)}{0.130}t \qquad \bar{R}^2 = 0.9852$$

(*) تفسير المتغيرات موجود في مثال 6.18 (الأخطاء القياسية موجودة بين الأقواس)

المصدر: G.S. Maddala, Econometrics, Mc Craw- Hill, New York, 1977, p. 242.

مثال 3.20:

نموذج تسعير مدخر رأس المال في شكل نظام تكراري

The Capital Asset pricing model expressed as a recursire system

في تطبيق غير معتاد لنماذج المعادلات الآنية قدر كل من Cheng F.Lee و W.P.Liayd[19] النموذج التالي لصناعة الزيوت.

$$R_{1t} = \alpha_1 + \gamma_1 M_t + u_{1t}$$

$$R_{2t} = \alpha_2 + \beta_{21}R_{1t} + \gamma_2 M_t + u_{2t}$$

$$R_{3t} = \alpha_3 + \beta_{31}R_{1t} + \beta_{32}R_{2t} + \gamma_3 M_t + u_{3t}$$

$$R_{4t} = \alpha_4 + \beta_{41}R_{1t} + \beta_{42}R_{2t} + \beta_{43}R_{3t} + \gamma_4 M_t + u_{4t}$$

$$R_{5t} = \alpha_5 + \beta_{51}R_{1t} + \beta_{52}R_{2t} + \beta_{53}R_{3t} + \beta_{54}R_{4t} + \gamma_5 M_t + u_{5t}$$

$$R_{6t} = \alpha_6 + \beta_{61}R_{1t} + \beta_{62}R_{2t} + \beta_{63}R_{3t} + \beta_{64}R_{4t} + \beta_{65}R_{5t} + \gamma_6 M_t + u_{6t}$$

$$R_{7t} = \alpha_7 + \beta_{71}R_{1t} + \beta_{72}R_{2t} + \beta_{73}R_{3t} + \beta_{74}R_{4t} + \beta_{75}R_{5t} + \beta_{76}R_{6t} + \gamma_7 M_t + u_{7t}$$

(19) The Capital Asset pricing model expressed as a recursive system: An Empirical Investigation, Journal of Financial and quantitative analysis, Jan 1976, pp.237- 249.

حيث إن :

R_1 = معدل العائد على الحماية 1 (= زيت imperial)

R_2 = معدل العائد على الحماية 2 (= زيت عباد الشمس)

⋮

R_7 = معدل العائد على الحماية 7 (= القياسي في Indiana)

M_t = معدل العائد على مؤشر السوق

u_{it} = مقادير الأخطاء (i = 1 , 2 , . . . ,7)

قبل استعراض النتائج، دعنا نطرح السؤال التالي: كيف نختار مستوى الحماية 1 أو مستوى الحماية 2 وهكذا؟ أجاب Lee و Lioyd عن هذا السؤال بشكل عملي. فقاما بعمل انحدار لمعدل العائد على الحماية i على معدل العائد على الحماية للمستويات الستة الباقية ويسجلها قيمة R^2. وبالتالي، فإن لدينا سبعة انحدارات من هذا الشكل، ثم قاما بترتيب قيم R^2 المقدرة من الأصغر للأكبر. مستوى الحماية صاحب أقل قيمة لـ R^2 يرمز له بمستوى الحماية 1 والمستوى الذي له أكبر قيمة لـ R^2 يرمز له بمستوى الحماية 7. الفكرة وراء ذلك بسيطة نسبيًا. إذا كانت R^2 لمعدل العائد الخاص بزيت imperial هي أصغر قيمة مقارنة مع مستويات الحماية الستة الباقية، فإن ذلك يقترح أن هذا المستوى من الحماية يتأثر أقل بالتغيير في عوائد مستويات الحماية الأخرى، وبالتالي الترتيب السببي، إذا وجد، يتحرك في الاتجاه من هذا المستوى إلى المستويات الأخرى وليس بالعكس.

على الرغم من ذلك، يعتبر تطبيقًا ضعيفًا للترتيب السببي، إلا أننا نعرض النتائج العملية في جدول (7.20).

في تمرين 5.5 نتعرض للشكل المميز لنظرية الاستثمار الحديثة، والتي ببساطة عبارة عن انحدار لمعدل العائد على الحماية I على معدل العائد على السوق. معامل الميل، المعروف باسم معامل Beta، يعتبر مقياسًا للتطاير على عائد الحماية. نتائج انحدار Lee- Lioyed تقترح وجود معنوية في العلاقات الموجودة بين عوائد الحماية وهذه المعنوية بعيدة عن أثر السوق المشترك الممثل في حقيبة السوق. وبالتالي فعائد Indiana القياسي يعتمد ليس فقط على معدل العائد في السوق ولكن أيضًا على معدل العائد لزيوت Shell، بترول Phillips وزيت Union. للتعبير عن ذلك بشكل مختلف، فإن التحركات في معدل العائد للمستوى القياسي لـ Indiana يمكن تفسيره بشكل أفضل، إذا وضعنا في الاعتبار معدل العائد لكل من زيوت Shell، بترول Phillips وزيوت Union بالإضافة إلى معدل العائد في السوق.

جدول (7.20) تقديرات نظام

	Linear form dependent variables						
	Standard of Indiana	Shell Oil	Phillips Petroleum	Union Oil	Standard of Ohio	Sun Oil	Imperial Oil
Standard of Indiana							
Shell Oil	0.2100* (2.859)						
Phillips Petroleum	0.2293* (2.176)	0.0791 (1.065)					
Union Oil	0.1754* (2.472)	0.2171* (3.177)	0.2225* (2.337)				
Standard of Ohio	−0.0794 (−1.294)	0.0147 (0.235)	0.4248* (5.501)	0.1468* (1.735)			
Sun Oil	0.1249 (1.343)	0.1710* (1.843)	0.0472 (0.355)	0.1339 (0.908)	0.0499 (0.271)		
Imperial Oil	−0.1077 (−1.412)	0.0526 (0.6804)	0.0354 (0.319)	0.1580 (1.290)	−0.2541* (−1.691)	0.0828 (0.971)	
Constant	0.0868 (0.681)	−0.0384 (1.296)	−0.0127 (−0.068)	−0.2034 (0.986)	0.3009 (1.204)	0.2013 (1.399)	0.3710* (2.161)
Market index	0.3681* (2.165)	0.4997* (3.039)	0.2884 (1.232)	0.7609* (3.069)	0.9089* (3.094)	0.7161* (4.783)	0.6432* (3.774)
R^2	0.5020	0.4658	0.4106	0.2532	0.0985	0.2404	0.1247
Durbin–Watson	2.1083	2.4714	2.2306	2.3468	2.2181	2.3109	1.9592

مثال 4.20 :

الشكل المصحح لنموذج St. Louis[20]

نموذج St. Louis الشهير والمثير للجدل، والذي تم عمله في أواخر 1960 تم تصحيحه من وقت لآخر. إحدى هذه المراجعات معطاة في جدول (8.20) (لاحظ أن: وجود نقط فوق المتغير تعني معدل نمو هذا المتغير). النموذج يشتمل أساسًا على المعادلات (1)، (2)، (4) و (5) في جدول (8.20)، المعادلات الأخرى تمثل التعريفات. معادلة (1) تم تقديرها باستخدام OLS.

جدول (8.20) نموذج St. Louis

(1) $$\dot{Y}_1 = C1 + \sum_{i=0}^{4} CM_i(\dot{M}_{t-i}) + \sum_{i=0}^{4} CE(\dot{E}_{t-i}) + \varepsilon 1_t$$

(2) $$\dot{P}_t = C2 + \sum_{i=1}^{4} CPE_i(\dot{PE}_{t-i}) + \sum_{i=0}^{5} CD_i(\dot{X}_{t-i} - \dot{XF}^*_{t-i1})$$
$$+ CPA(\dot{PA}_t) + CDUM1(DUM1) + CDUM2(DUM2) + \varepsilon 2_t$$

(3) $$\dot{PA}_t = \sum_{i=1}^{21} CPRL_i(\dot{P}_{t-i})$$

(20) Review, Federal Reserve Bank of St. Louis. May 1982, p.14.

$$(4)\quad RL_t = C3 + \sum_{i=0}^{20} CPRL_i(\dot{P}_{t-i}) + \varepsilon 3_t$$

$$(5)\quad U_t - UF_t = CG(GAP_t) + CG1(GAP_{t-1}) + \varepsilon 4_t$$

$$(6)\quad Y_t = (P_t/100)(X_t)$$

$$(7)\quad \dot{Y}_t = [(Y_t/Y_{t-i})^4 - 1]100$$

$$(8)\quad \dot{X}_t = [(X_t/X_{t-i})^4 - 1]100$$

$$(9)\quad \dot{P}_t = [(P_t/P_{t-i})^4 - 1]100$$

$$(10)\quad GAP_t = [(XF_t/X_t)/XF_t]100$$

$$(11)\quad \dot{XF}_t^* = [(XF_t/X_{t-1})^4 - 1]100$$

Y = GUP الاسمي
M = مخزون المال (M1)
E = مصاريف العمالة المرتفعة
P = GNP المنكمش (1972=100)
PE = السعر النسبي للطاقة
X = الناتج في 1972 دولار
XF = الناتج المحتمل (Rasche/ Tatom)
RL = معدل السندات المشترك
U = معدل البطالة
UF = معدل البطالة في العمالة الكاملة
DUN 1 = متغير وهمي تحكمي (III–1971 إلى I–1973=1 ، 0 بخلاف ذلك)
DUN 2 = متغير وهمي للتحكم البعيد (II–1973 إلى I–1975 =1 ، 0 بخلاف ذلك)

المصدر: Federal reserve bank of St. Louis, review, May, 1983. p.14

المعادلات (1)، (2) و (4) تم تقديرها باستخدام طريقة Almon للتوزيعات المتأخرة، والتي تشتمل على قيود (نقطة النهاية) للمعاملات. في حالة الضرورة تم تصحيح بعض المعادلات من الارتباط التسلسلي من الدرجة الأولى (ρ_1) أو من الدرجة الثانية (ρ_2) أو كليهما.

بتحليل النتائج، نرى أن معدل النمو في المعروض من المال يحدد معدل النمو الإسمي في GNP وليس معدل النمو في مصاريف العمالة العالية. مجموع معاملات M يساوي 1.06 مما يعني أن 1% زيادة محتملة في متوسط المعروض من المال يؤدي إلى حوالي 1.06% زيادة في الـ GNP الإسمي. على الجانب الآخر مجموع معاملات الـ E حوالي 0.05 مما يعني أن التغيير في مصاريف العمالة العالية له تأثير صغير على معدل النمو في الـ GNP الإسمي. متروك للقارئ تفسير نتائج الانحدارات الأخرى الموجودة في جدول (9.20).

جدول (9.20) تقديرات العينة : I-1960 إلى IV-1980 (القيم المطلقة لإحصاء t موجودة بين الأقواس)

(1)
$$\hat{Y}_t = \underset{(2.15)}{2.44} + \underset{(3.38)}{0.40\dot{M}_t} + \underset{(5.06)}{0.39\dot{M}_{t-1}} + \underset{(2.18)}{0.22\dot{M}_{t-2}} + \underset{(0.82)}{0.06\dot{M}_{t-3}} - \underset{(0.11)}{0.01\dot{M}_{t-4}}$$
$$+ \underset{(1.46)}{0.06\dot{E}_t} + \underset{(0.63)}{0.02\dot{E}_{t-1}} - \underset{(0.57)}{0.02\dot{E}_{t-2}} - \underset{(0.52)}{0.02\dot{E}_{t-3}} + \underset{(0.34)}{0.01\dot{E}_{t-4}}$$
$$R^2 = 0.39 \quad se = 3.50 \quad DW = 2.02$$

(2)
$$\widehat{\dot{P}_t} = \underset{(2.53)}{0.96} + \underset{(0.75)}{0.01\dot{PE}_{t-1}} + \underset{(1.96)}{0.04\dot{PE}_{t-2}} - \underset{(0.73)}{0.01\dot{PE}_{t-3}} + \underset{(1.38)}{0.02\dot{PE}_{t-4}}$$
$$- \underset{(0.18)}{0.00(\dot{X}_t - \dot{XF}^*_t)} + \underset{(1.43)}{0.01(\dot{X}_{t-1} - \dot{XF}^*_{t-1})} + \underset{(4.63)}{0.02(\dot{X}_{t-2} - \dot{XF}^*_{t-2})}$$
$$+ \underset{(3.00)}{0.02(\dot{X}_{t-3} - \dot{XF}^*_{t-3})} + \underset{(2.42)}{0.02(\dot{X}_{t-4} - \dot{XF}^*_{t-4})} + \underset{(2.16)}{0.01(\dot{X}_{t-5} - \dot{XF}^*_{t-5})}$$
$$+ \underset{(10.49)}{1.03(\dot{PA}_t)} - \underset{(1.02)}{0.61(DUM1_t)} + \underset{(2.71)}{1.65(DUM2_t)}$$
$$R^2 = 0.80 \quad se = 1.28 \quad DW = 1.97 \quad \hat{\rho} = 0.12$$

(4)
$$\widehat{RL_t} = \underset{(3.12)}{2.97} + \underset{(5.22)}{0.96}\sum_{i=0}^{20} \dot{P}_{t-i}$$
$$R^2 = 0.32 \quad se = 0.33 \quad DW = 1.76 \quad \hat{\rho} = 0.94$$

(5)
$$\widehat{U_t - UF_t} = \underset{(11.89)}{0.28(GAP_t)} + \underset{(6.31)}{0.14(GAP_{t-1})}$$
$$R^2 = 0.63 \quad se = 0.17 \quad DW = 1.95 \quad \hat{\rho}_1 = 1.43 \quad \hat{\rho}_2 = 0.52$$

المصدر: Federal reserve bank of St. Louis, review, May, 1983. p.14

7.20 الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

1 - بافتراض أن معادلة ما داخل نموذج معادلات آنية موصفة (سواء تامة التوصيف أو موصفة بأكثر مما يجب) فإن لدينا طرقًا عديدة لتقديرها.

2 - هذه الطرق تقسم إلى فئتين رئيسيتين: طرق المعادلة المنفردة، وطرق الأنظمة.

3 - لأسباب اقتصادية أو أخطاء في التعريف، أو ما غير ذلك من أسباب أخرى، تعتبر طرق المعادلة المنفردة هي الأكثر استخدامًا. صفة مميزة في مثل هذه الطرق ألا وهي إمكانية تقدير معادلة منفردة ما في نموذج متعدد المعادلات بدون الاهتمام بالمعادلات الأخرى الموجودة في النظام. (لاحظ أنه: بغرض التوصيف فإنه يجب الوضع في الاعتبار المعادلات الأخرى).

4 - هناك ثلاث طرق للمعادلة المنفردة تستخدم بكثرة هذه الطرق هي: OLS و ILS و 2SLS.

5 – على الرغم من أن طريقة OLS في العموم غير مناسبة للاستخدام في إطار نماذج المعادلات الآنية، إلا أنه يمكن تطبيقها فيما يسمى النماذج، حيث توجد علاقة سبب وآثار واضحة وذات اتجاه واحد فقط بين المتغيرات الداخلية.

6 – طريقة ILS مناسبة أكثر للمعادلات تامة التوصيف. في هذه الطريقة يتم تطبيق OLS على المعادلة المخفضة الشكل، ومن المعادلات المخفضة يتم تقدير المعاملات البنائية الأصلية.

7 – طريقة 2SLS مصممة خصيصًا للمعادلات الموصفة بأكثر مما يجب، وعلى الرغم من ذلك يمكن تطبيقها على المعادلات تامة التوصيف، ولكن في مثل هذه الحالة تتطابق نتائج 2SLS مع نتائج ILS الفكرة الرئيسية وراء 2SLS هي استبدال المتغير المفسر الداخلي (العشوائي) بتوليفة خطية من المتغيرات المحددة سابقًا في النموذج، ونستخدم هذه التوليفة كمتغير مفسر بدلاً من المتغير الداخلي الأصلي. وبالتالي فإن طريقة 2SLS تتشابه مع طريقة المتغيرات المساهمة في التقدير، حيث إن التوليفة الخطية من المتغيرات المحددة سابقًا تمثل المتغير المساهم، أو المفوض للمتغير المنحدر الداخلي.

8 – تتميز مقدرات كل من ILS و 2SLS بأنها مقدرات متسقة، بمعنى أنه مع زيادة حجم العينة فإن التقديرات تؤول إلى قيم المجتمع الحقيقية.

هذه المقدرات قد لاتحقق صفات العينات صغيرة الحجم، مثل عدم التحيز والتباين الأقل. وبالتالي فالنتائج التي يتم الحصول عليها من هذه المقدرات في العينات صغيرة الحجم والاستدلال الإحصائي المعتمد على مثل هذه التقديرات، يجب تفسيره بمزيد من الحرص والدقة.

تمارين : EXERCISES

أسئلة : Questions

1.20 حدد أيًا من العبارات التالية صح أو خطأ:

(a) طريقة OLS غير مناسبة لتقدير المعادلة البنائية في نموذج المعادلات الآنية.

(b) في حالة المعادلات غير الموصفة، 2SLS تكون غير مناسبة للتطبيق.

(c) مشكلة الآنية غير موجودة في نموذج المعادلات الآنية ذي الاتجاه الواحد.

(d) مشكلة الآنية ومشكلة خارجية المنشأ تعني نفس الشئ.

(e) طريقة 2SLS وبعض الطرق الأخرى لتقدير المعادلات البنائية لها خصائص إحصائية جيدة فقط في حالة العينات كبيرة الحجم.

(f) لا توجد قيمة مماثلة لـ R^2 في حالة نموذج المعادلات الآنية ككل.

(g) (*)- طريقة 2SLS وبعض الطرق الأخرى لتقدير المعادلات البنائية تكون غير مناسبة إذا كان الخطأ في المعادلة مرتبطًا ذاتيًا أو مرتبطًا مع نظيره في المعادلات الأخرى أو كليهما.

(h) إذا كانت المعادلة تامة التوصيف، فإن ILS و 2SLS تعطي نفس النتائج.

2.20 لماذا من غير الضروري تطبيق طريقة المربعات الصغرى ذات المرحلتين على المعادلات تامة التوصيف؟

3.20 اعتبر نموذج Keynesian المعدل التالي لتحديد الدخل:

$$C_t = \beta_{10} + \beta_{11} Y_t + u_{1t}$$
$$I_t = \beta_{20} + \beta_{21} Y_t + \beta_{22} Y_{t-1} + u_{2t}$$
$$Y_t = C_t + I_t + G_t$$

حيث إن:

C = نفقات الاستهلاك

I = نفقات الاستثمار

Y = الدخل

G = نفقات حكومية

G_t و Y_{t-1} مفترض أنهما محددان سابقًا.

(a) احصل على المعادلات المخفضة الشكل، وحدد أيًا من المعادلات السابقة موصف (سواء تام التوصيف أو موصف بأكثر مما يجب).

(b) أي الطرق ستستخدم لتقدير معلمات المعادلة الموصفة بأكثر مما يجب والمعادلة تامة التوصيف؟ علل إجابتك.

(*) اختياري.

4.20 اعتبر النتائج التالية(*):

$$OLS: \quad \widehat{\dot{W}_t} = 0.276 + 0.258\dot{P}_t + 0.046\dot{P}_{t-1} + 4.959V_t$$

$$OLS: \quad \widehat{\dot{P}_t} = 2.693 + 0.232\dot{W}_t - 0.544\dot{X}_t + 0.247\dot{M}_t + 0.064\dot{M}_{t-1}$$

$$2SLS: \quad \widehat{\dot{W}_t} = 0.272 + 0.257\dot{P}_t + 0.046\dot{P}_{t-1} + 4.966V_t$$

$$2SLS: \quad \widehat{\dot{P}_t} = 2.686 + 0.233\dot{W}_t - 0.544\dot{X}_t + 0.246\dot{M}_t + 0.046\dot{M}_{t-1}$$

حيث إن $\dot{W}_t$ ، $\dot{P}_t$ ، $\dot{M}_t$ و $\dot{X}_t$ هي نسب التغير في الكسب، الأسعار، أسعار السلع المستوردة وإنتاجية العمالة (تغيرات النسب خلال السنة السابقة) بالترتيب. أما V_t فتمثل التعطل عن العمل بسبب الإجازات (نسب إجمالي العمالة).

"بما أن نتائج OLS و 2SLS متماثلة تمامًا، فإنه لايوجد معنى لاستخدام طريقة 2SLS". علق على ذلك.

5.20(†) افترض أن الإنتاج يمكن وصفه من خلال دالة الإنتاج لـ Cobb-Douglas :

$$Q_i = AK_i^{\alpha} L_i^{\beta}$$

بحيث إن:

Q = النتائج

K = وحدة رأس المال

T = وحدة العمالة

A، α و β = معلمات

i = المصنع رقم i

بمعلومية سعر الناتج النهائي P، سعر العمالة W، وسعر رأس المال R، وبافتراض تعظيم الربح نحصل على هذا النموذج العملي للإنتاج:

دالة الإنتاج:

$$\ln Q_i = \ln A + \alpha \ln K_i + \beta \ln L_i + \ln u_{1i} \qquad (1)$$

(*) المصدر:

Prices and Earnings in 1951–1969 : An Econometrics Assessment, Department of Employment, United Kngdam, Her Majesty's Stationery Office, London, 1971, p. 30.

(†) اختياري.

الناتج الحدي لدالة العمالة:

$$\ln Q_i = -\ln\beta + \ln L_i + \ln\frac{W}{P} + \ln u_{2i} \qquad (2)$$

الناتج الحدي لدالة رأس المال:

$$\ln Q_i = -\ln\alpha + \ln K_i + \ln\frac{R}{P} + \ln u_{3i} \qquad (3)$$

حيث إن u_1، u_2 و u_3 تمثل الأخطاء العشوائية.

في النموذج السابق، هناك ثلاث معادلات في ثلاثة متغيرات داخلية Q، L و K أما المتغيرات K، P، R و W تعتبر متغيرات خارجية.

(a) ما هي المشاكل التي ستواجهك عند تقدير النموذج إذا كان $\alpha + \beta = 1$ أي عندما يكون هناك ثابت في عائد المقياس؟

(b) إذا كان $\alpha + \beta \neq 1$، هل يمكنك تقدير هذه المعادلات؟ ضع في الاعتبار مسألة إمكانية توصيف النظام.

(c) إذا كان النظام لايمكن توصيفه. ما الذي يمكن عمله لجعله قابلاً للتوصيف؟ لاحظ أن: المعادلتين (2) و (3) تم الحصول عليهما من تفاضل Q بالنسبة للعمالة و رأس المال بالترتيب، ثم مساواة ذلك مع W/P و R/P تم تحويل النتيجة في صورة لوغاريتم وإضافة مقدار الخطأ.

6.20 اعتبر نموذج الطلب- العرض للمال التالي:

الطلب على المال: $M_t^d = \beta_0 + \beta_1 Y_1 + \beta_2 R_t + \beta_3 P_t + u_{1t}$

العرض على المال: $M_t^s = \alpha_0 + \alpha_1 Y_t + u_{2t}$

بحيث إن: M = المال

Y = الدخل

R = معدل الفائدة

P = السعر

افترض أن R و P محددان من قبل (متغيرات محددة سابقًا).

(a) هل دالة الطلب يمكن توصيفها؟

(b) هل دالة العرض يمكن توصيفها؟

(c) أي الطرق يمكن استخدامها لتقدير المعادلة أو المعادلات التي يمكن توصيفها؟ لماذا؟

(d) افترض أننا عدلنا دالة العرض بإضافة المتغيرات المفسرة Y_{t-1} و M_{t-1} ماذا سيحدث لمشكلة التوصيف؟ هل ستظل تستخدم نفس الطريقة في التقديرات والتي استخدمتها في c؟ علل إجابتك.

7.20 بالعودة إلى تمرين 10.18. للنظام ثنائي المعادلات، احصل على المعادلات مخفضة الشكل، وقدر معلمات النظام. قدر انحدار المربعات الصغرى غير المباشرة للاستهلاك على الدخل، وقارن هذه النتائج مع نظيرها الخاص بانحدار الـ OLS.

مسائل : Problems

8.20 اعتبر النموذج التالي:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + u_{1t}$$
$$Y_t = \alpha_0 + \alpha_1 R_t + u_{2t}$$

حيث M_t (عرض المال) متغير خارجي، R_t معدل الفائدة و Y_t هو GDP.

(a) كيف يمكنك تفسير النموذج؟

(b) هل المعادلات يمكن توصيفها؟

(c) باستخدام البيانات المعطاة في جدول (2.20). قدر معلمات المعادلات القابلة للتوصيف. علل اختيارك لهذا النموذج.

9.20 افترض أننا عدلنا النموذج الموجود في تمرين 8.20 كالتالي:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + \beta_3 Y_{t-1} + u_{1t}$$
$$Y_t = \alpha_0 + \alpha_1 R_t + u_{2t}$$

(a) هل هذا النظام يمكن توصيفه؟

(b) باستخدام البيانات المعطاة في جدول (2.20)، قدر معلمات المعادلات القابلة للتوصيف.

10.20 اعتبر النموذج التالي:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + u_{1t}$$
$$Y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + u_{2t}$$

حيث هذه المتغيرات معرفة كما هو موجود في تمرين 8.20. اعتبر I (الاستثمار المحلي) و M متغيرات خارجية. حدد إمكانية توصيف النظام. باستخدام بيانات جدول (2.20)، قدر معلمات المعادلة أو المعادلات القابلة للتوصيف.

11.20 افترض أننا عدلنا النموذج الموجود في تمرين 10.20 كالتالي:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + u_{1t}$$
$$Y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + u_{2t}$$
$$I_t = \gamma_0 + \gamma_1 R_t + u_{3t}$$

افترض أن M متغير خارجي محدد سابقًا.

(a) حدد أي المعادلات يمكن توصيفها.

(b) قدر معلمات المعادلة أو المعادلات القابلة للتوصيف باستخدام البيانات المعطاة في جدول (2.20). علل اختيارك لطرق التقدير التي استخدمتها.

12.20 – اثبت الأخطاء القياسية الموجودة في (3.5.20).

13.20 بالعودة إلى نموذج الطلب- العرض المعطى في المعادلتين (1.3.20) و (2.3.20). افترض أن دالة العرض معرفة كالتالي:

$$Q_t = \beta_0 + \beta_1 P_{t-1} + u_{2t}$$

حيث P_{t-1} هو السعر السائد في الفترة السابقة.

(a) إذا كانت X (النفقات) و P_{t-1} متغيرات محددة سابقًا. هل توجد مشكلة الآنية؟

(b) إذا كانت هذه المشكلة موجودة، هل يمكن توصيف دوال العرض والطلب؟ إذا كان من الممكن توصيفها، احصل على المعادلات المخفضة الشكل وقدرها من البيانات المعطاة في جدول (1.20).

(c) من المعاملات المخفضة، هل يمكن استنتاج المعاملات البنائية؟ وضح الحسابات الضرورية لذلك.

14.20 تمرين في الفصل: اعتبر نموذج الاقتصاد الكلي المبسط التالي لاقتصاد الولايات المتحدة، مثلاً في الفترة من 1960–1999[*] . دالة الاستهلاك الخاص:

دالة الاستثمار الخاص:

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + u_{1t} \qquad \alpha_1 > 0, 0 < \alpha_2 < 1$$

دالة الاستثمار الكلي الخاص:

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 I_{t-1} + u_{2t} \qquad \beta_1 > 0, \beta_2 < 0, 0 < \beta_3 < 1$$

دالة الطلب على المال:

$$R_t = \lambda_0 + \lambda_1 Y_t + \lambda_2 M_{t-1} + \lambda_3 P_t + \lambda_4 R_{t-1} + u_{3t}$$
$$\lambda_1 > 0, \lambda_2 < 0, \lambda_3 > 0, 0 < \lambda_4 < 1$$

تعريف الدخل:

$$Y_t = C_t + I_t + G_t$$

حيث C = الاستهلاك الخاص الحقيقي، I = الاستثمار الكلي الخاص الحقيقي، G = النفقات الحكومية، GDP=Y الحقيقي، M = عرض المال M_2 عند الأسعار الحالية، R = معدل الفائدة للمدى البعيد (%) و P = مؤشر سعر المستهلك. المتغيرات الداخلية هي C، I، R و Y. المتغيرات المحددة سابقًا هي: C_{t-1}، I_{t-1}، M_{t-1}، P_t، R_{t-1} و G_t بالإضافة إلى الجزء المقطوع من المحور الصادي. الـ u's هي مقادير الأخطاء.

(a) باستخدام شرط الترتيب للتوصيف. حدد أيًا من المعادلات الأربع قابلة للتوصيف، سواء تام التوصيف أو موصف بأكثر مما يجب.

(b) أي الطرق ستستخدم لتقدير المعادلة أو المعادلات القابلة للتوصيف؟

(c) احصل على بيانات مناسبة سواء حكومية أو خاصة أو كلاهما معًا لتقدير النموذج، وعلق على نتائجك.

(*) مأخوذ من:
K. A. Lawler, and A. V. Katos Econometrics: A Practical Approach, Routledge, New York, 2000, p. 204.

APPENDIX 20A ملحق A 20

1.A 20 التحيز في مقدرات المربعات الصغرى غير المباشرة : BAIS IN THE INDIRECT LEAST- SQUARES ESTIMATORS

لتوضيح أن مقدرات ILS، على الرغم من اتساقها، مقدرات متحيزة، قمنا باستخدام نموذج الطلب- العرض المعطى في المعادلتين (1.3.20) و (2.3.20). من المعادلة (10.3.20) نحصل على:

$$\hat{\beta}_1 = \frac{\hat{\Pi}_3}{\hat{\Pi}_1}$$

والآن من (7.3.20)

$$\hat{\Pi}_3 = \frac{\sum q_t x_t}{\sum x_t^2}$$

نجد أن :

ومن (5.3.20) نجد أن:

$$\hat{\Pi}_3 = \frac{\sum q_t x_t}{\sum x_t^2}$$

وبالتالي بالتعويض نحصل على :

$$\hat{\beta}_1 = \frac{\sum q_t x_t}{\sum p_t x_t} \tag{1}$$

باستخدام (3.3.20) و (4.3.20) نحصل على:

$$p_t = \Pi_1 x_t + (w_t - \bar{w}) \tag{2}$$

$$q_t = \Pi_3 x_t + (v_t - \bar{v}) \tag{3}$$

حيث إن $\bar{w}$ و $\bar{v}$ هما القيم المتوسطة لـ w_t، v_t بالترتيب.

بالتعويض عن (2) و (3) في (1) نحصل على:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\Pi_3 \sum x_t^2 + \sum (v_t - \bar{v}) x_t}{\Pi_1 \sum x_t^2 + \sum (w_t - \bar{w}) x_t} \\ &= \frac{\Pi_3 + \sum (v_t - \bar{v}) x_t / \sum x_t^2}{\Pi_1 + \sum (w_t - \bar{w}) x_t / \sum x_t^2}\end{aligned} \tag{4}$$

بما أن معامل التوقع E هو معامل خطي، لاتستطيع الحصول على توقع (4)، على الرغم من وضع أن $\hat{\beta}_1 \neq \Pi_3/\Pi_1$ بوجه عام. (لماذا؟) ولكن مع زيادة حجم

العينة إلى مالانهاية، نحصل على :

$$\text{plim}(\hat{\beta}_1) = \frac{\text{plim}\,\Pi_3 + \text{plim}\,\sum(v_t - \bar{v})x_t / \sum x_t^2}{\text{plim}\,\Pi_1 + \text{plim}\,\sum(w_t - \bar{w})x_t / \sum x_t^2} \qquad (5)$$

وقد تم استخدام العينة التالية لـ Plim وهي :

$$\text{plim}(A + B) = \text{plim}\,A + \text{plim}\,B$$

و

$$\text{plim}\left(\frac{A}{B}\right) = \frac{\text{plim}\,A}{\text{plim}\,B}$$

والآن مع زيادة حجم العينة، فإن المقدر الثاني في كل من البسط والمقام في (5) سيؤول إلى الصفر (لماذا؟) وسيؤدي ذلك إلى :

$$\text{plim}(\hat{\beta}_1) = \frac{\Pi_3}{\Pi_1} \qquad (6)$$

مما يوضح أنه بالرغم من التحيز، فإن $\hat{\beta}_1$ مقدار متسق لـ β_1.

2.A 20 تقدير الأخطاء القياسية لمقدرات 2SLS :
ESTIMATION OF STANDARD ERRORS OF 2SLS ESTIMATORS

الهدف من هذا الملحق، هو إثبات أن الأخطاء القياسية للمقدرات التي يتم الحصول عليها من الانحدار الثاني في طريقة 2SLS، باستخدم المعادلة التطبيقية في مقدرات OLS ليس التقدير "المناسب" للأخطاء القياسية "الحقيقية". لتوضيح ذلك، دعنا نستخدم نموذج الدخل - عرض المال المعطى في (1.4.20) و (2.4.20). قدرنا معلمات دالة عرض المال الموصفة بأكثر مما يجب من انحدار المرحلة الثانية كالتالي :

$$Y_{2t} = \beta_{20} + \beta_{21}\hat{Y}_{1t} + u_t^* \qquad (6.4.20)$$

حيث :

$$u_t^* = u_{2t} + \beta_{21}\hat{u}_t \qquad (7)$$

والآن عندما نقوم بعمل الانحدار (6.4.20)، الأخطاء القياسية لـ $\hat{\beta}_{21}$ مثلاً يتم الحصول عليها من الشكل التالي :

$$\hat{\sigma}_{u^*}^2 = \frac{\sum(\hat{u}_t^*)^2}{n-2} = \frac{\sum(Y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}\hat{Y}_{1t})^2}{n-2} \qquad (8)$$

حيث :

$$\hat{\sigma}_{u^*}^2 = \frac{\sum(\hat{u}_t^*)^2}{n-2} = \frac{\sum(Y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}\hat{Y}_{1t})^2}{n-2} \quad (9)$$

ولكن σ_u^2 ليس مساويًا تمامًا لـ $\hat{\sigma}_{u_2}^2$ ، حيث إن الأخير مقدر غير متحيز للتباين الحقيقي u_2. الفرق يمكن إثباته من (7). للحصول على $\hat{\sigma}_{u_2}^2$ الحقيقي (كما عرف سابقًا)، نقوم بالتالي :

$$\hat{u}_{2t} = Y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}Y_{1t}$$

حيث $\hat{\beta}_{20}$ و $\hat{\beta}_{21}$ هي مقدرات انحدار المرحلة الثانية وبالتالي :

$$\hat{\sigma}_{u_2}^2 = \frac{\sum(Y_{2t} - \hat{\beta}_{20} - \hat{\beta}_{21}Y_{1t})^2}{n-2} \quad (10)$$

لاحظ الفرق بين (9) و (10) : في (10) نستخدم Y_1 الحقيقية بدلاً من تقدير Y_1 من انحدار المرحلة الأولى.

بعد تقدير (10)، أفضل طريقة لتصحيح الأخطاء القياسية للمعاملات المقدرة في انحدار المرحلة الثانية، يتم بضرب كل واحدة منها في $\hat{\sigma}_{u_2} / \hat{\sigma}_{u*}$ لاحظ أنه إذا كانت قيم Y_{1t} و $\hat{Y}_{1t}$ قريبة من بعضها البعض، بمعنى أن R^2 في انحدار المرحلة الأولى تكون قيمة عالية جداً، فإن معامل التصحيح $\hat{\sigma}_{u_2} / \hat{\sigma}_{u*}$ سيكون قريبًا من 1، في مثل هذه الحالة، فإن الأخطاء القياسية المقدرة في انحدار المرحلة الثانية يمكن اعتبارها تقديرات حقيقية. ولكن في المواقف الأخرى، سنحتاج إلى معامل التصحيح السابق ذكره.

الفصل الحادي والعشرون

السلاسل الزمنية في الاقتصاد القياسي: بعض المفاهيم الأساسية

TIME SERIES ECONOMETRICS: SOME BASIC CONCEPTS

كما سبق وذكرنا في الفصل (1)، فإن بيانات السلاسل الزمنية تعتبر واحدة من أهم أنواع البيانات المستخدمة في التحليل العملي. في هذه المقدمة، وفي الفصل التالي، سنقترب أكثر إلى مثل هذا النوع من البيانات، ليس فقط بسبب كثرة استخدامها في الواقع، ولكن أيضًا بسبب التحديات العديدة التي تطرحها للباحثين وعلماء الاقتصاد القياسي.

أولاً: التجارب العملية المعتمدة على بيانات سلاسل زمنية، تفترض أن السلسلة الزمنية محل الاهتمام ساكنة. على الرغم من أننا ناقشنا مفهوم السكون في الفصل 1، إلا أننا سنناقشه كاملاً في هذا الفصل. وبشكل أكثر تحديداً، سنحاول بالضبط تحديد ما يعنيه السكون، ولماذا يجب وضعه في الاعتبار عند الدراسة.

ثانيًا: في الفصل (12)، الخاص بالارتباط الذاتي، ناقشنا أسبابًا عديدة للارتباط الخطي. أحيانًا يكون الارتباط الذاتي نتيجة لأن السلسلة الزمنية محل الاهتمام غير ساكنة.

ثالثًا : في انحدار متغير سلسلة زمنية على متغير أو متغيرات أخرى من نوع السلاسل الزمنية غالبًا ما يحصل الفرد على R^2 ذات قيمة عالية جدًا (تزيد عن 0.9) حتى ولو كان لا توجد علاقة لها معنى بين المتغيرات. فأحيانًا نحن نتوقع عدم وجود أي علاقة بين متغيرين حتى ولو كان انحدار أحدهما على الآخر يظهر علاقة معنوية. هذا الموقف يعبر عن مشكلة الانحدار الزائف أو غير ذي معنى، والذي ستتم مناقشة

طبيعته لاحقًا. وبالتالي، فإنه من المهم معرفة ما إذا كانت العلاقة بين المتغيرات الاقتصادية زائفة أو ليست لها معنى أم لا. سنرى في هذا الفصل، كيف يظهر الانحدار الزائف إذا كانت السلسلة الزمنية غير ساكنة.

رابعًا: بعض السلاسل الزمنية المالية، مثل أسعار الأسهم، يظهر فيها ما يسمى بظاهرة السير العشوائي. وذلك يعني أفضل تنبؤ لسعر السهم مثلاً IBM، غداً يساوي سعره اليوم بالإضافة إلى جزء عشوائي صاف (أو مقدار خطأ). إذا كان ذلك هو الحل، فإن تقدير أسعار الممتلكات سيكون عملاً عقيمًا غير ذي جدوى.

خامسًا: نماذج الانحدار التي تشتمل على بيانات سلاسل زمنية، غالبًا ما تستخدم بغرض التنبؤ. وبالنظر إلى المناقشة السابقة، نجد أنه من الضروري معرفة ما إذا كان هذا التنبؤ سليمًا، إذا كانت السلسلة الزمنية محل الاهتمام غير ساكنة أم لا.

أخيراً: اختبارات السببية لكل من Granger و Sims والتي ناقشناها في الفصل (17)، تفترض أن السلسلة الزمنية الموجودة في التحليل ساكنة. وبالتالي فاختبار سكون السلسلة الزمنية لابد أن يسبق اختبار السببية.

موضوع تحليل السلاسل الزمنية يعتبر من المواضيع المهمة العديدة التفاصيل، والخلفية الرياضية الخاصة بتحليل السلاسل الزمنية معقدة ومتداخلة، بحيث إن أفضل ما نستطيع تقديمه في كتاب مبادئ عامة مثل هذا الكتاب، هو بدايات المفاهيم الأساسية لتحليل السلاسل الزمنية. ولمن يرد معرفة المزيد من التفاصيل عن هذا الموضوع، عليه النظر في المراجع المعطاة والخاصة بذلك[1].

(1) كمستوى تعريفي أو أساسي للموضع، المراجع المفيدة في ذلك هي:
At the introductory level these references may be helpful: Gary Koop Analysis of Economic Data John Wiley & Sons New York 2000, Jeff B. Cromwell, Walter C. Labys and Michel Terraza, Univariate Tests for Time Series Models, Sage Publications, Califomia, Ansbury Park, 1994; Jeff B. Cromwell, Michael H. Hannan, Walter C. Labys, and Michel Terraza, Mulhvariate Tests for Tme Series Models, Sage Publications, Califomia, Ansbury Park, 1994; H. R. Seddighi, K. A. Lawler, and A. V. Katos, Economertics: A Practical Approach, Routledge, New York, 2000. At the intermediate level, see Walter Enders, Applied Econometric Tme Series, John Wiley & Sons, New York, 1995; Kerry Patterson, An Introductton to Applied Econometrics: A Tme Series Approach St. Martin's Press New York 2000 T. C. Mills The Econometric ModelAng of Financial nme Series 2d ed., Cambridge University Press New York 1999 Mamo Verbeek A Guide to Modem Econometncs, John Wiley & Sons, New York, 2000; Wojciech W. Charemza and Derek F. Deadman, New Dtrections in Econometric Fractice: General to Specil'ic Modelling and VectorAutoregression, 2d ed., Edward Elgar Publisher, New York, 1997. At the advanced level, see Hamilton, J. D., Tme Series Analysis, Princeton University Press, Princeton, N.J., 1994, and G. S. Madda a and In-Moo Kim, Unit Roots, Cointegration, and Stnuctural Change, Cambridge University Press, 1998. At the applied level, see B. Bhaskara Rao, ed., Cointegration for the Applied Economist St. Martin's Press New York 1994 and Chandan Mukherjee Howard White and Marc Wuyts, Econometrics and Data Analysis for Developing Countries, Routledge, New York, 1998.

1.21 نظرة على بعض السلاسل الزمنية الاقتصادية المختارة للولايات المتحدة: A LOOK AT SELECTED U.S. ECONOMIC TIME SERIES

لبداية عرض الموضوع، وتقديم فكرة مبسطة عن تحليل السلاسل الزمنية للقارئ في هذا الفصل، قد يكون من المفيد اعتبار سلاسل زمنية اقتصادية للولايات المتحدة في مواضيع متعددة. السلاسل الزمنية التي سيتم اعتبارها هنا هي: (1) GDP (الناتج الكلي المحلي)، (2) PDI (الدخل الشخصي المتغير)، (3) PCE (نفقات الاستهلاك الشخصي)، (4) الربح (الربح بعد الضرائب) و (5) الأرباح الموزعة على المساهمين (الأرباح الصافية الموزعة على المساهمين المشتركة)، كل البيانات مسجلة ببلايين الدولارات في سنة 1987، وللفترات الربع سنوية في المدة من 1970–1971 بإجمالي 88 مشاهدة ربع سنوية. البيانات الأصلية معطاة في جدول (1.21).

جدول (1.21) بيانات اقتصاد كلي ، الولايات المتحدة ، I–1970 إلى IV–1991 .

Quarter	GDP	PDI	PCE	Profits	Dividend	Quarter	GDP	PDI	PCE	Profits	Dividend
1970–I	2872.8	1990.6	1800.5	44.7	24.5	1981–I	3860.5	2783.7	2475.5	159.5	64.0
1970–II	2860.3	2020.1	1807.5	44.4	23.9	1981–II	3844.4	2776.7	2476.1	143.7	68.4
1970–III	2896.6	2045.3	1824.7	44.9	23.3	1981–III	3864.5	2814.1	2487.4	147.6	71.9
1970–IV	2873.7	2045.2	1821.2	42.1	23.1	1981–IV	3803.1	2808.8	2468.6	140.3	72.4
1971–I	2942.9	2073.9	1849.9	48.8	23.8	1982–I	3756.1	2795.0	2484.0	114.4	70.0
1971–II	2947.4	2098.0	1863.5	50.7	23.7	1982–II	3771.1	2824.8	2488.9	114.0	68.4
1971–III	2966.0	2106.6	1876.9	54.2	23.8	1982–III	3754.4	2829.0	2502.5	114.6	69.2
1971–IV	2980.8	2121.1	1904.6	55.7	23.7	1982–IV	3759.6	2832.6	2539.3	109.9	72.5
1972–I	3037.3	2129.7	1929.3	59.4	25.0	1983–I	3783.5	2843.6	2556.5	113.6	77.0
1972–II	3089.7	2149.1	1963.3	60.1	25.5	1983–II	3886.5	2867.0	2604.0	133.0	80.5
1972–III	3125.8	2193.9	1989.1	62.8	26.1	1983–III	3944.4	2903.0	2639.0	145.7	83.1
1972–IV	3175.5	2272.0	2032.1	68.3	26.5	1983–IV	4012.1	2960.6	2678.2	141.6	84.2
1973–I	3253.3	2300.7	2063.9	79.1	27.0	1984–I	4089.5	3033.2	2703.8	155.1	83.3
1973–II	3267.6	2315.2	2062.0	81.2	27.8	1984–II	4144.0	3065.9	2741.1	152.6	82.2
1973–III	3264.3	2337.9	2073.7	81.3	28.3	1984–III	4166.4	3102.7	2754.6	141.8	81.7
1973–IV	3289.1	2382.7	2067.4	85.0	29.4	1984–IV	4194.2	3118.5	2784.8	136.3	83.4
1974–I	3259.4	2334.7	2050.8	89.0	29.8	1985–I	4221.8	3123.6	2824.9	125.2	87.2
1974–II	3267.6	2304.5	2059.0	91.2	30.4	1985–II	4254.8	3189.6	2849.7	124.8	90.8
1974–III	3239.1	2315.0	2065.5	97.1	30.9	1985–III	4309.0	3156.5	2893.3	129.8	94.1
1974–IV	3226.4	2313.7	2039.9	86.8	30.5	1985–IV	4333.5	3178.7	2895.3	134.2	97.4
1975–I	3154.0	2282.5	2051.8	75.8	30.0	1986–I	4390.5	3227.5	2922.4	109.2	105.1
1975–II	3190.4	2390.3	2086.9	81.0	29.7	1986–II	4387.7	3281.4	2947.9	106.0	110.7
1975–III	3249.9	2354.4	2114.4	97.8	30.1	1986–III	4412.6	3272.6	2993.7	111.0	112.3
1975–IV	3292.5	2389.4	2137.0	103.4	30.6	1986–IV	4427.1	3266.2	3012.5	119.2	111.0
1976–I	3356.7	2424.5	2179.3	108.4	32.6	1987–I	4460.0	3295.2	3011.5	140.2	108.0
1976–II	3369.2	2434.9	2194.7	109.2	35.0	1987–II	4515.3	3241.7	3046.8	157.9	105.5
1976–III	3381.0	2444.7	2213.0	110.0	36.6	1987–III	4559.3	3285.7	3075.8	169.1	105.1
1976–IV	3416.3	2459.5	2242.0	110.3	38.3	1987–IV	4625.5	3335.8	3074.6	176.0	106.3
1977–I	3466.4	2463.0	2271.3	121.5	39.2	1988–I	4655.3	3380.1	3128.2	195.5	109.6
1977–II	3525.0	2490.3	2280.8	129.7	40.0	1988–II	4704.8	3386.3	3147.8	207.2	113.3
1977–III	3574.4	2541.0	2302.6	135.1	41.4	1988–III	4734.5	3407.5	3170.6	213.4	117.5
1977–IV	3567.2	2556.2	2331.6	134.8	42.4	1988–IV	4779.7	3443.1	3202.9	226.0	121.0
1978–I	3591.8	2587.3	2347.1	137.5	43.5	1989–I	4809.8	3473.9	3200.9	221.3	124.6
1978–II	3707.0	2631.9	2394.0	154.0	44.5	1989–II	4832.4	3450.9	3208.6	206.2	127.1
1978–III	3735.6	2653.2	2404.5	158.0	46.6	1989–III	4845.6	3466.9	3241.1	195.7	129.1
1978–IV	3779.6	2680.9	2421.6	167.8	48.9	1989–IV	4859.7	3493.0	3241.6	203.0	130.7
1979–I	3780.8	2699.2	2437.9	168.2	50.5	1990–1	4880.8	3531.4	3258.8	199.1	132.3

1979–II	3784.3	2697.6	2435.4	174.1	51.8	1990–II	4900.3	3545.3	3258.6	193.7	132.5
1979–III	3807.5	2715.3	2454.7	178.1	52.7	1990–III	4903.3	3547.0	3281.2	196.3	133.8
1979–IV	3814.6	2728.1	2465.4	173.4	54.5	1990–IV	4855.1	3529.5	3251.8	199.0	136.2
1980–I	3830.8	2742.9	2464.6	174.3	57.6	1991–I	4824.0	3514.8	3241.1	189.7	137.8
1980–II	3732.6	2692.0	2414.2	144.5	58.7	1991–II	4840.7	3537.4	3252.4	182.7	136.7
1980–III	3733.5	2722.5	2440.3	151.0	59.3	1991–III	4862.7	3539.9	3271.2	189.6	138.1
1980–IV	3808.5	2777.0	2469.2	154.6	60.5	1991–IV	4868.0	3547.5	3271.1	190.3	138.5

لاحظ أن: GDP (الناتج الكلي المحلي)، ببلايين الدولارات 1987، P.A- 96.

PDI (الدخل الشخصي المتغير)، بلايين الدولارات 1987، P.A- 112.

PCE (نفقات الاستهلاك الشخصي)، بلايين الدولارات 1987، P.A- 96.

الأرباح (الأرباح المشتركة بعد الضرائب)، بلايين الدولارات، P.A- 110.

الأرباح الموزعة على المساهمين (الأرباح الصافية الموزعة على المساهمين)، بلايين الدولارات، P.A- 110.

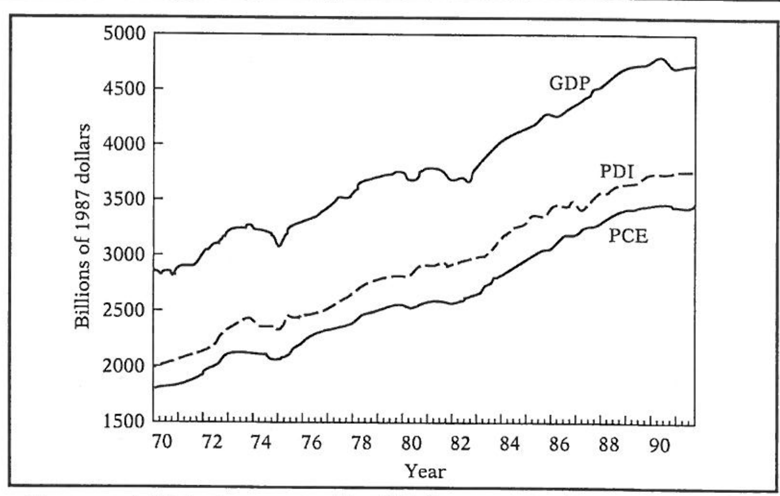
المصـــدر: U.S. Department of commerce, Bureau of Economic Analysis, Business Statistics, 1963-1991, June 1992

شكل (1.21) يعبر عن بيانات الـ GDP، PDI و PCE، وشكل (2.21) يمثل السلسلتين الزمنيتين الأخريين. الرسم البياني للبيانات يعتبر دائمًا الخطوة الأولى في تحليل أي سلسلة زمنية. الانطباع الأول الذي يتكون من هذه الرسوم البيانية، أن كل السلاسل الزمنية الموضحة في الشكلين (1.21) و (2.21) يظهر عليها الاتجاه لأعلى، مع وجود بعض التموجات (التقلبات).

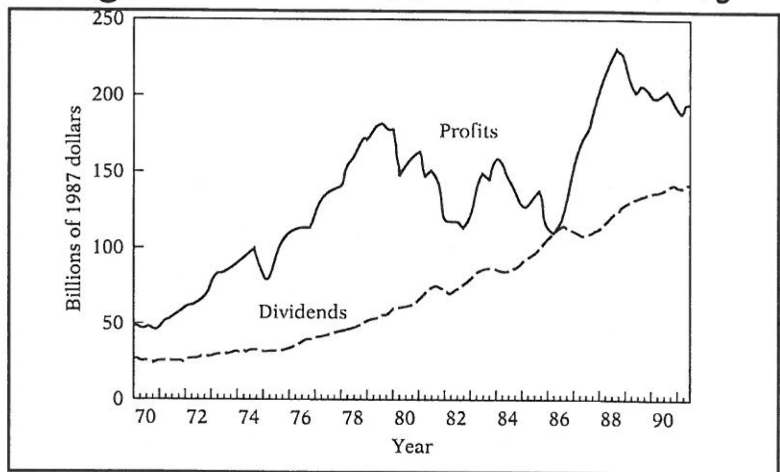
افترض أننا نريد التنبؤ بشكل هذه المنحنيات عبر فترات ربع سنوية، مثلاً من I–1992 إلى 1996–IV[2]. هل نستطيع ببساطة تكملة المنحنيات الموجودة في هذه الأشكال؟ يمكن ذلك إذا عرفنا الطريقة العشوائية أو الإحصائية التي تسير بها البيانات، أو إذا عرفنا كيف تتم عملية توليد البيانات، والتي تم بها عمل هذه المنحنيات؟ ولكن ما هي هذه الطريقة؟

الإجابة على ذلك وأسئلة أخرى مرتبطة بنفس الموضوع، نحتاج إلى معرفة بعض المصطلحات "الجديدة" والتي صممت خصيصًا لتحليل السلاسل الزمنية والتي سننتقل إليها الآن.

(2) بالطبع فإن لدينا البيانات الحقيقية لهذه الفترة الآن، ونستطيع مقارنة هذه البيانات الحقيقية مع نظيرها الذي تم التنبؤ به على أساس البيانات في فترات سابقة.



شكل (1.21) GDP ، PDI و PCE ، الولايات المتحدة ، 1970– 1991 (ربع سنوية)



شكل (2.21) الأرباح والأرباح الموزعة على المساهمين ، الولايات المتحدة 1970–1991 (ربع سنوية)

2.21 مفاهيم أساسية [3]: KEY CONCEPTS

ماهي هذه المصطلحات؟ هي تشمل التالي :

1 – عمليات عشوائية .

2 – عمليات ساكنة .

3 – عمليات تامة العشوائية .

(3) المناقشة التالية مأخوذة من

MADDALA ETAL., OP.CIT., AND CHAREMZA ETAL., OP.CIT.

4 – عمليات غير ساكنة.
5 – متغيرات الدمج.
6 – نماذج السير العشوائي.
7 – الاندماج المزدوج.
8 – الاتجاهات العامة المحددة والعشوائية.
9 – اختبارات جذر الوحدة

فيما يلي، سنناقش كلاً من هذه المفاهيم. مناقشتنا ستكون غالبًا مشجعة لمزيد من الاكتشاف. وسيتم تقديم بعض الأمثلة المناسبة عندما يكون ذلك ممكنًا وضروريًا.

3.21 العمليات العشوائية : STOCHASTIC PROCESSES

العملية العشوائية هي مجموعة من المتغيرات مرتبة زمنيًا(4). إذا جعلنا Y ترمز لمتغير عشوائي، وإذا كان متغيرًا عشوائيًا متصلاً، نرمز له بالرمز $Y(t)$، ولكن إذا كان متغيرًا متقطعًا، نرمز له بالرمز Y_t. وتعتبر الصورة البيانية الكهربائية للقلب مثالاً للمتغير الأخير، ومثالاً للمتغير الأول هو GDP، PDI وهكذا. وبما أن معظم البيانات الاقتصادية مجموعة في نقاط متقطعة من الزمن، سنستخدم الرمز Y_t بدلاً من $Y(t)$، فإذا كانت Y تمثل GDP، ففي بياناتنا يكون لدينا Y_1، Y_2، Y_3، Y_4، . . . ، Y_{86}، Y_{87}، Y_{88}، حيث إن الترميز 1 يرمز إلى المفردة الأولى (بمعنى GDP في الربع الأول من 1970) و الترميز 88 يرمز إلى آخر مفردة (بمعنى GDP في الربع الرابع من 1991). ضع في الاعتبار أن كلاً من هذه الـ Y's هي متغيرات عشوائية.

كيف إذن يمكن النظر إلى GDP على أنه عملية عشوائية؟ اعتبر على سبيل المثال الـ GDP المساوي لـ 2872.8 بليون دولار لـ 1970–I. في النظرية فإن قيمة GDP في الربع الأول من 1970 قد يكون أي قيمة، معتمدًا على الوضع السياسي والاقتصادي السائد. القيمة 2872.8 يتم فيها التعبير عن كل ذلك(5). وبالتالي يمكن القول بأن الـ GDP

(4) كلمة " عشوائي " تأتي أساسًا من كلمة يونانية "STOKHOS" والتي تعني الهدف إذا رميت مرة أسهمًا على لوحة تنشين الأسهم بغرض إصابة النقطة السوداء المتوسطة بها كم مرة ستحقق ذلك؟ من مئات الأسهم يمكن أن تكون محظوظا لإصابة الهدف بعض المرات، وفي المرات الأخرى الأسهم ستكون موزعة عشوائيًا حول نقطة الهدف.

(5) يمكن النظر إلى القيمة 2872.8$ كقيمة متوسطة لكل القيم الممكنة لـ GDP في الربع الأول من 1970.

عملية عشوائية والقيم الحقيقية التي نشاهدها للفترة I–1970 إلى IV–1991 هي تمثيل محدد لهذه العملية (بمعنى عينة) التفرقة بين العملية العشوائية وتمثيلها مماثل للتفرقة بين المجتمع والعينة في بيانات Cross- section. فكما تستخدم بيانات العينة للوصول إلى استدلالات عن المجتمع، ففي السلاسل الزمنية تستخدم التمثيل للوصول إلى استدلالات عن العملية العشوائية محل الاهتمام.

عمليات عشوائية ساكنة: Stationary Stochastic Processes

إحدي العمليات العشوائية التي حصلت على اهتمام كبير في مجال تحليل السلاسل الزمنية هي العملية المسماة بالعملية العشوائية الساكنة. بوجه عام، فإن العملية العشوائية يقال عنها ساكنة إذا كان المتوسط والتباين ثابتين بمرور الزمن، وقيمة التغاير بين فترتين زمنيتين تعتمد فقط على المسافة أو الفجوة أو الفترة الزمنية المتأخرة بين هاتين الفترتين وليس على الوقت الحقيقي الذي يحسب عنده التغاير. في تاريخ السلاسل الزمنية مثل هذه العملية العشوائية تعرف باسم السكون الضعيف، أو سكون التغاير، أو السكون من الدرجة الثانية، أو عملية عشوائية بمعنى واسع. في هذا الفصل، ومعظم المواقف التطبيقية يكون هذا النوع من العمليات العشوائية كافيًا ووافيًا بالغرض[6].

لشرح السكون الضعيف، دع Y_t تمثل سلسلة زمنية عشوائية بهذه الخصائص:

(1.3.21) **الوسط**: $E(Y_t) = \mu$

(2.3.21) **التباين**: $\operatorname{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$

(3.3.21) **التغاير**: $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

حيث γ_k، التغاير (أو التغاير الذاتي) عند الفترة الزمنية المتأخرة k، هو التغاير بين القيم Y_t و Y_{t+k}، بمعنى أنه بين قيمتين للـ Y بينهما فترة زمنية k. إذا كانت $k = 0$ نحصل على γ_0 والذي يمثل ببساطة تباين Y (σ^2)، وإذا كانت $k = 1$، فإن γ_1 يمثل التغاير بين قيمتين متجاورتين للـ Y وهذا هو نوع التغاير الذي تعرضنا له في الفصل (12) (تذكر طريقة Markov للانحدار الذاتي من الدرجة الأولى).

---

(6) العملية العشوائية تكون تامة السكون إذا كانت كل عزوم التوزيع الاحتمالي وليس فقط أول اثنين (المتوسط والتباين) غير متغيرين مع الزمن. إذا كانت عمومًا عملية السكون تتبع التوزيع الطبيعي، فإن العملية العشوائية ذات السكون الضعيف تكون أيضًا تامة السكون، حيث إن العملية العشوائية التابعة للتوزيع الطبيعي تكون معرفة بالكامل من خلال أول عزمين، الوسط والتباين.

افترض أننا نقلنا نقط الأصل لـ Y من Y_t إلى Y_{t+m} (مثلاً من الربع الأول في 1970 إلى الربع الأول في 1975 لبيانات الـ GDP). الآن إذا كان Y_t ساكنًا فإن المتوسط، التباين والتغاير الذاتي لـ Y_{t+m} لابد أن يتساويا مع نظيرهما لـ Y_t. باختصار إذا كانت السلسلة الزمنية ساكنة، فإن توقعها وتباينها و التغاير الذاتي لها (عند فترات زمنية متأخرة عديدة) تظل كما هي بغض النظر عن النقطة الزمنية التي يتم قياسها فيها، بمعنى أنها لا تتغير مع الزمن. مثل هذا النوع من السلاسل الزمنية سيتجه أكثر للعودة إلى متوسطه (ويسمى ذلك العودة للمتوسط) والانقلاب حول هذا المتوسط (ويتم قياس ذلك بالتباين) سيكون له شكل ثابت نوعًا ما[7].

إذا كانت السلسلة الزمنية غير ساكنة بالمعنى السابق ذكره، تسمى بالسلسلة الزمنية غير الساكنة (تذكر أننا نتحدث فقط عن السكون الضعيف). بمعنى آخر، السلسلة الزمنية غير الساكنة سيكون لديها متوسط متغير مع الزمن، أو تباين متغير مع الزمن أو كليهما.

لماذا يعتبر سكون السلسلة الزمنية أمرًا في غاية الأهمية؟ لأنه إذا كانت السلسلة الزمنية غير ساكنة، فنحن نستطيع دراسة سلوكها فقط في الفترة الزمنية محل الدراسة. فكل مجموعة من بيانات السلاسل الزمنية ستكون خاصة فقط بالمرحلة محل الاهتمام. وكنتيجة لذلك، لايكون من الممكن تعميم ما نحصل عليه من نتائج على فترات زمنية أخرى. وبالتالي، فإن مثل هذا النوع من السلاسل الزمنية ستكون له قيمة طبيعية صغيرة خصوصًا لتحقيق هدف التنبؤ.

كيف يمكننا معرفة إذا كانت السلسلة الزمنية ساكنة أم لا؟ في الواقع، هل السلاسل الزمنية المعطاة في شكلي (1.21 و 2.21) ساكنة؟ سندرس هذه النقطة في الفقرة 8.21 و 9.21 حيث سنقوم باستعراض العديد من الاختبارات الخاصة بالسكون. ولكن إذا اعتمدنا على الحس المشترك فإنه يمكن القول بأن السلاسل الزمنية الموجودة في شكلي (1.21 و 2.21) غير ساكنة، على الأقل في القيم المتوسطة. ولكن المزيد من تلك التفاصيل سيتم استعراضها لاحقًا.

قبل الانتقال إلى نقطة أخرى، فقد ذكرنا نوعًا خاصًا من العمليات العشوائية (أو السلاسل الزمنية) والمسمى بالعشوائية الخالصة أو white noise process عملية تغيرات عشوائية بحتة.

---

(7) هذه النقطة تم طرحها في:

Keith Cutlibertson, Stephen G. Hall And Mark P. Taylor, Applied Econometric Techniques, The University Of Michigan Press, 1995, P.130.

سنسمي العملية العشوائية بأنها خالصة العشوائية إذا كان متوسطها يساوي صفرًا وله تباين ثابت σ^2 وغير مرتبطة تسلسليًا[8]. ودعنا نتذكر مقدار الخطأ u_t والموجود في نموذج الانحدار الخطي ذي فروض التوزيع الطبيعي التقليدية، والتي ناقشناها في الجزء I من هذا الكتاب، والتي تفترض أن هذا الخطأ يعتبر white noise process خطأ عملية تغيرات عشوائية بحتة، والتي نرمز لها بالرمز $u_t \sim \text{IIDN}(0, \sigma^2)$ بمعنى أن u_t مستقلة وموزعة تماثليًا كتوزيع طبيعي له متوسط صفر وتباين ثابت.

العمليات العشوائية غير الساكنة: Nonstationary Stochastic Processes

على الرغم من أننا نهتم أكثر بالعمليات العشوائية الساكنة، فإن الفرد أحيانًا يتعرض إلى سلسلة زمنية غير ساكنة، والمثال التقليدي على ذلك، هو نموذج السير العشوائي (RWM)[9]. وغالبًا ما يقال إن أسعار الممتلكات مثل أسعار الأسهم أو معدل تغير العملات يتبع سيرًا عشوائيًا، بمعنى أنه غير ساكن، وسنفرق بين نوعين من السير العشوائي: (1) سير عشوائي بدون اتجاه (بمعنى عدم وجود مقدار ثابت، أو الجزء المقطوع من المحور الصادي غير موجود) و (2) سير عشوائي مع الاتجاه (بمعنى أن المقدار الثابت موجود).

السير العشوائي بدون الاتجاه: Random Walk Without Drift

افترض أن u_t هي مقدار خطأ عملية تغيرات عشوائية بحتة له توقع 0 وتباين σ^2. وبالتالي فالسلسلة Y_t يقال إنها تمثل سيرًا عشوائيًا إذا كان:

$$Y_t = Y_{t-1} + u_t \qquad \text{(4.3.21)}$$

في نموذج السير العشوائي، كما في (4.3.21) توضح أن قيمة Y عند الزمن t تتساوى مع قيمتها عند الزمن $(t-1)$ بالإضافة إلى صدمة أو خطأ عشوائي، بمعنى أنها تتبع نموذج AR(1) باستخدام مصطلحات الفصلين (12 و 17). يمكن التفكير في (4.3.21) كمعادلة انحدار لـ Y عند الزمن t على قيمة لـ Y أيضًا، ولكن في فترة زمنية متأخرة واحدة. المؤيدون لكفاءة فرض سوق المال يعتقدون أن أسعار الأسهم هي

(8) إذا كانت أيضًا مستقلة فإنها تسمى عملية عشوائية بحتة.

(9) مصطلح السير العشوائي غالبًا مايقارن بسير الفرد السكران. عندما يترك البار ويتحرك مسافة عشوائية u_t عند الزمن t ويكمل السير بشكل غير متزن وينجرف بعيدًا عن البار. يقال نفس الشيء عن أسعار الأسهم. فاليوم سعر السهم يساوي سعر السهم بالأمس بالإضافة إلى صدمة عشوائية.

بالضرورة عشوائية، وبالتالي لايوجد مجال للتنبؤ الدقيق في سوق الأسهم، فإذا استطاع أحد التنبؤ بالسعر غداً بناء على سعر اليوم، فهو بالتأكيد مليونير. الآن من (4.3.21) يمكن كتابة التالي :

$$Y_1 = Y_0 + u_1$$
$$Y_2 = Y_1 + u_2 = Y_0 + u_1 + u_2$$
$$Y_3 = Y_2 + u_3 = Y_0 + u_1 + u_2 + u_3$$

وعمومًا إذا بدأت العملية عند وقت ما 0 بقيمة لـ Y_0، سيكون لدينا

$$Y_t = Y_0 + \sum u_t \qquad \textbf{(5.3.21)}$$

وبالتالي :

$$E(Y_t) = E\left(Y_0 + \sum u_t\right) = Y_0 \qquad \text{(لماذا)} \qquad \textbf{(6.3.21)}$$

وبنفس الطريقة ، يمكن إيضاح أن

$$\text{var}\,(Y_t) = t\sigma^2 \qquad \textbf{(7.3.21)}$$

كما يتضح مما سبق، فإن متوسط Y يساوي قيمتها الأصلية أو قيمتها عند نقطة البداية، والتي تعتبر ثابتًا، ولكن كلما زادت t فإن التباين يزداد، وبالتالي فذلك يتعارض مع شرط السكون. باختصار فإن RWM بدون اتجاه هي عملية عشوائية غير ساكنة. في الواقع Y_0 تساوي الصفر، وفي مثل هذه الحالة يكون $E\,(Y_t) = 0$.

صفة مميزة لـ RWM هي ثبات الصدمات العشوائية (بمعنى الأخطاء العشوائية)، ويتضح ذلك من (5.3.21): Y_t هي مجموع القيمة الأصلية Y_0، بالإضافة إلى مجموع الصدمات العشوائية. كنتيجة لذلك، فإن أثر صدمة معينة لاينتهي بمرور الوقت. فمثلاً، إذا كان $u_2 = 2$ بدلاً من $u_2 = 0$ فإن كل الـ Y_t's بداية من Y_2 سيكون أكبر بوحدتين، وأثر هذه الصدمة لن ينتهي، ولهذا السبب فإن السير العشوائي يقال إن لديه ذاكرة لا نهائية، كما لاحظ Kerry Patterson فإن السير العشوائي يتذكر كالصدمات للأبد(10). أي أن له ذاكرة لانهائية.

ومثير للاهتمام أنه إذا كتبنا (4.3.21) كالتالي :

$$(Y_t - Y_{t-1}) = \Delta Y_1 = u_t \qquad \textbf{(8.3.21)}$$

حيث إن Δ هي معامل الفرق الأول الذي ناقشناه من قبل في الفصل (12). من السهل إثبات أنه عندما تكون Y_t غير ساكنة، فإن الفرق الأول لها يكون ساكنًا، بمعنى

(10) Kerry Patterson, Op Cit., Chap.6.

آخر الفرق الأول لسلسلة زمنية لسير عشوائي تعتبر سلسلة ساكنة، ولدينا المزيد من هذه النقطة، ولكن سنتناولها لاحقًا.

السير العشوائي باتجاه. دعنا نعدل (4.3.21) كالتالي:

$$Y_t = \delta + Y_{t-1} + u_1 \qquad \textbf{(9.3.21)}$$

حيث إن δ معروفة كمعامل الاتجاه. اسم الاتجاه يأتي من حقيقة أنه عندما نكتب المعادلة السابقة كالتالي:

$$Y_t - Y_{t-1} = \Delta Y_t = \delta + u_t \qquad \textbf{(10.3.21)}$$

فإنها توضح اتجاه Y_t سواء كان لأعلى أو لأسفل، معتمدًا على δ كونها موجبة أو سالبة. لاحظ أن النموذج (9.3.21) هو أيضًا نموذج AR(1). وباتباع الطريقة المذكورة سابقًا في السير العشوائي بدون اتجاه، من الممكن إثبات أن نموذج السير العشوائي باتجاه (9.3.21) له التالي:

$$E(Y_t) = Y_0 + t \cdot \delta \qquad \textbf{(11.3.21)}$$

$$\text{var}(Y_t) = t\sigma^2 \qquad \textbf{(12.3.21)}$$

كما ترى، فبالنسبة لـ RWM باتجاه، فإن المتوسط والتباين أيضًا يتزايدان مع الزمن، مما يعتبر مرة أخرى مخالفة لشرط السكون (الضعيف). باختصار فإن RWM سواء باتجاه أو بدون اتجاه هي عملية عشوائية غير ساكنة.

لتوضيح السير العشوائي باتجاه وبدون اتجاه، قمنا بعمل المحاكاتين التاليتين:

$$Y_t = Y_0 + u_t \qquad \textbf{(13.3.21)}$$

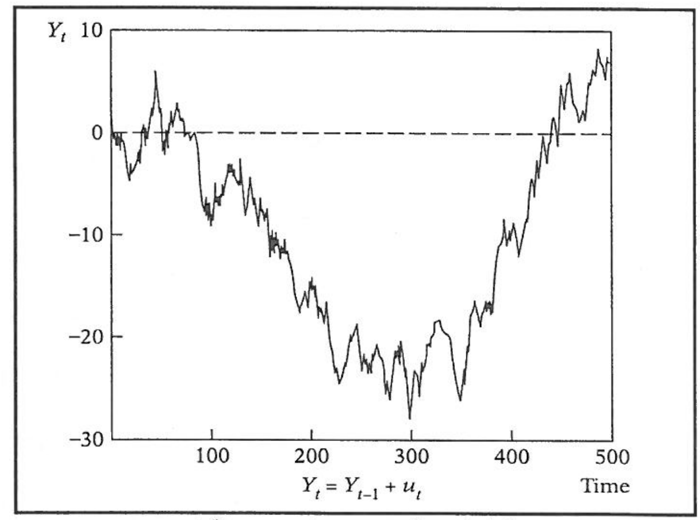
حيث إن u_t هو مقدار الخطأ العشوائي البحت، بمعنى أن كل $u_t \sim N(0, 1)$ أي أن u_t تتبع التوزيع الطبيعي القياسي. من مولد الأرقام العشوائية نحصل على 500 قيمة لـ u ونولد Y_t كما هو موضح في (13.3.21). افترضنا أن $Y_0 = 0$. وبالتالي (13.3.21) هو RWM بدون اتجاه.

الآن اعتبر التالي:

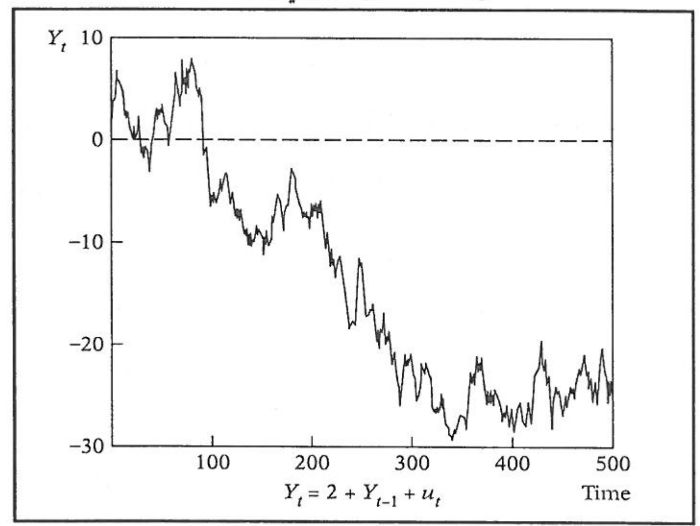
$$Y_t = \delta + Y_0 + u_t \qquad \textbf{(14.3.21)}$$

والذي يمثل RWM باتجاه. افترضنا أن u_t و Y_0 كما في (13.3.21) وافترضنا أن $\delta = 2$. الرسم البياني لنماذج (13.3.21) و (14.3.21) بالترتيب معطاة في الشكلين (3.21 و 4.21). يمكن للقارئ أن يقارن الشكلين في ضوء مناقشتنا لـ RWM باتجاه وبدون اتجاه.

نموذج سير العشوائي هو مثال لما هو معروف سابقًا باسم عملية جذور الوحدة. وبما أن هذا المصطلح له أهمية كبيرة في تاريخ السلاسل الزمنية، فإننا سنقوم بتفسير ماهية عملية جذور الوحدة.



شكل (3.21) سير عشوائي بدون اتجاه



شكل (4.21) سير عشوائي باتجاه

4.21 عملية جذر الوحدة العشوائي :
UNIT ROOT STOCHASTIC PROCESS

دعنا نكتب RWM (4.3.21) كالتالي :

$$Y_t = \rho Y_{t-1} + u_t \qquad -1 \leq \rho \leq 1 \qquad (1.4.21)$$

هذا النموذج يفترض نموذج انحدار ذاتي من الدرجة الأولى لـ Morkov والذي ناقشناه في فصل الارتباط الذاتي. إذا كانت $\rho = 1$، فإن (1.4.21) تصبح RWM (بدون اتجاه). إذا كانت ρ في الحقيقة تساوي 1، فإننا نواجه ما هو معروف باسم مشكلة جذور الوحدة، بمعنى أنه موقف عدم سكون، ونحن نعلم مسبقًا أنه في مثل هذه الحالة، يكون تباين Y_t غير ساكن. الاسم جذور الوحدة يأتي بسبب حقيقة أن $\rho = 1$ [11]. وبالتالي فالمصطلحات عدم السكون، السير العشوائي، وجذور الوحدة يمكن التعامل معها كمرادفات.

إذا كان عمومًا $|\rho| \leq 1$ فإن ذلك يحدث إذا كانت القيمة المطلقة لـ ρ أقل من الواحد، وبالتالي، من الممكن إثبات أن السلسلة الزمنية Y_t ساكنة بالمعنى الذي عرفناه[12].

في الواقع، من الضروري معرفة إذا كانت السلسلة الزمنية تعتبر عملية جذر وحدة أم لا[13]. في الفقرة 9.21 سنناقش اختبارات عديدة لجذور الوحدة، بمعنى اختبارات عديدة للسكون. في هذه الفقرة سنحدد أيضًا ما إذا كانت السلاسل الزمنية الموجودة في شكلي (1.21 و 2.21) ساكنة أم لا. ومحتمل أن يعتقد القارئ أن هذه السلاسل غير ساكنة، ولكن سنرى ما إذا كان ذلك صحيحًا أم لا.

---

(11) نقطة فنية: إذا كانت $\rho = 1$ فيمكن كتابة (1.4.21) كالتالي: $Y_t - Y_{t-1} = u_t$ والآن باستخدام معامل الفترات الزمنية المتأخرة L فإن: $LY_t = Y_{t-1}, L^2Y_t = Y_{t-2}$ وهكذا، يمكننا كتابة (1.4.21) كالتالي: $(1-L)Y_t = u_t$. المصطلح جذر الوحدة يشير إلى جذر متعدد الحدود في معامل الفترات الزمنية المتأخرة. إذا وضعنا $(1 - L) = 0$ نحصل على $L = 1$ ومن هنا جاءت التسمية جذر الوحدة.

(12) إذًا في (1.4.21) افترضنا أن القيمة المبدئية لـ $Y=(Y_0)$ تساوي الصفر، $|\rho| < 1$ و u_t هو خطأ عشوائي بحت يتبع التوزيع الطبيعي بتوقع يساوي الصفر وتباين الوحدة فإن $E(Y_t)=0$ و $\text{var}(Y_t)= 1/(1-\rho|^2)$. وبالتالي فإن كلاً منهما ثابت، وبالتعريف للسكون الضعيف فإن Y_t ساكنة. على الجانب الآخر، كما رأينا من قبل، إذا كانت $|\rho|=1$، Y_t هي سير عشوائي أو غير ساكنة.

(13) السلسلة الزمنية قد تحتوي على أكثر من جذر وحدة واحد. ولكن سنناقش هذا الوضع لاحقًا في هذا الفصل.

5.21 عمليات عشوائية ساكنة ذات اتجاه عامه وأخرى ذات فروق : TREND STATIONARY (TS) AND DIFFERENCE STATIONARY (DS) STOCHASTIC PROCESSES

التفرقة بين العملية العشوائية الساكنة وغير الساكنة (أو السلاسل الزمنية) لها حد حاسم خاص بما إذا كان الاتجاه العام (التطور طويل المدى البطئ للسلسلة الزمنية محل الاهتمام) الملاحظ في السلسلة الزمنية المتكونة في الشكلين (3.21 و 4.21) أو في السلسلة الزمنية الاقتصادية الحقيقية الموجودة في الشكلين (1.21 و 2.21) محدد أو عشوائي. بوجه عام إذا كان الاتجاه العام لسلسلة زمنية قابلاً للتنبؤ كاملاً وليس متغيراً سنسمي ذلك اتجاهًا عامًا محددًا. وبخلاف ذلك، إذا كان غير متنبأ به سنسميه اتجاهًا عامًا عشوائيًا. لوضع التعريف في شكل رسمي، دعنا نعتبر النموذج التالي للسلسلة الزمنية Y_t.

$$Y_1 = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + u_t \qquad (1.5.21)$$

حيث u_t هي مقدار الخطأ للعملية العشوائية البحتة، و t زمن مقاس بشكل فيه ترتيب زمني، والآن لدينا الحالات الممكنة التالية:

السير العشوائي البحت : Pure random walk

إذا كان في (1.5.21) $\beta_1 = 0$، $\beta_2 = 0$ و $\beta_3 = 1$ فإن لدينا:

$$Y_t = Y_{t-1} + u_t \qquad (2.5.21)$$

والتي ليست إلا RWM بدون اتجاه، وبالتالي غير ساكنة، ولكن لاحظ أنه إذا كتبنا (2.5.21) كالتالي:

$$\Delta Y_t = (Y_t + Y_{t-1}) = u_t \qquad (8.3.21)$$

ستصبح ساكنة كما لاحظنا من قبل، وبالتالي فإن RWM بدون اتجاه هي عملية ساكنة في الفرق (DSP).

السير العشوائي باتجاه : Random Walk with Drift

إذا كان في (1.5.21) $\beta_1 \neq 0$، $\beta_2 = 0$، $\beta_3 = 1$ فنحصل على:

$$Y_t = \beta_1 + Y_{t-1} + u_t \qquad (3.5.21)$$

والتي تمثل سيراً عشوائيًا باتجاه ، وبالتالي هي غير ساكنة ، وإذا كتبناها كالتالي :

$$(Y_t - Y_{t-1}) = \Delta Y_t = \beta_1 + u_t \qquad \text{(a.3.5.21)}$$

فإن هذا يعني أن Y_t لديها اتجاه عام موجب $(\beta_1 > 0)$ أو سالب $(\beta_1 < 0)$ (انظر شكل 4.21) ومثل هذا الاتجاه يسمى اتجاهًا عامًا عشوائيًا. المعادلة (a.3.5.21) تسمى عملية DSP، حيث إن عدم السكون الموجود في Y_t ممكن إلغاؤه بأخذ الفروق الأولى للسلسلة الزمنية.

الاتجاه العام المحدد :

إذا كان في (1.5.21)، $\beta_1 \neq 0$، $\beta_2 \neq 0$، $\beta_3 = 0$ فنحصل على :

$$Y_t = \beta_1 + \beta_{2t} + u_t \qquad \text{(4.5.21)}$$

والذي يسمى عملية ساكنة ذات اتجاه عام (TSP) على الرغم من أن متوسط Y_t يساوي $\beta_1 + \beta_{2t}$ والذي لايعتبر مقداراً ثابتًا وتباينه هو$(= \sigma^2)$ بمجرد معرفة قيم β_1، β_2، فإنه يمكن التنبؤ بقيم المتوسط بالضبط، وبالتالي إذا طرحنا متوسط Y_t من Y_t فإن السلسلة الناتجة ستكون ثابتة. من هنا جاءت التسمية سكون ذو اتجاه عام. هذه العملية الخاصة بحذف الاتجاه (المحدد) تسمى إزالة الاتجاه.

السير العشوائي باتجاه مع وجود اتجاه عام محدد :

Random Walk with Drift and deterministic trend

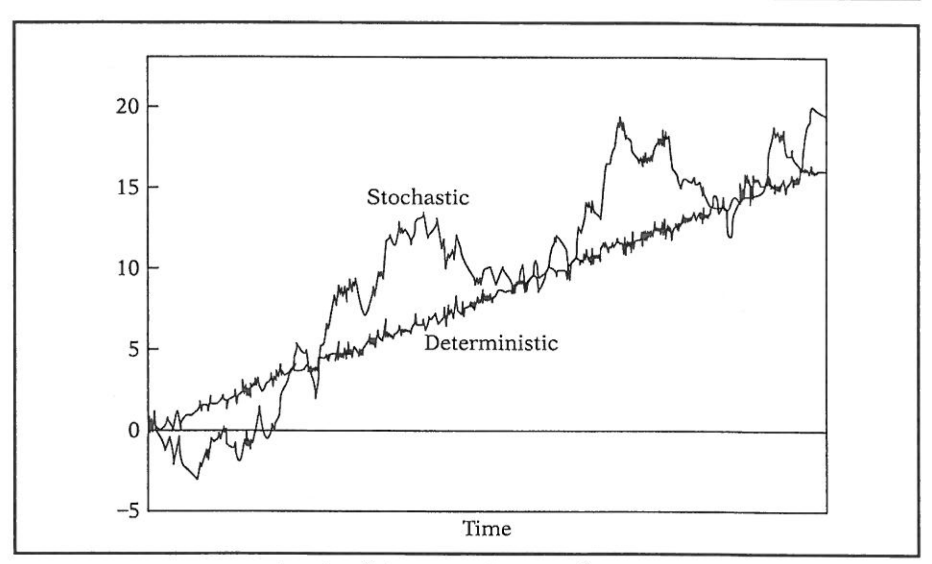
إذا كان في (1.5.21)، $\beta_1 \neq 0$ و $\beta_2 \neq 0$، $\beta_3 = 1$ نحصل على :

$$\mathrm{Y}_t = \beta_1 + \beta_{2t} + Y_{t-1} + u_t \qquad \text{(5.5.21)}$$

هنا لدينا سير عشوائي باتجاه، مع وجود اتجاه عام محدد، والذي يمكن رؤيته بوضوح إذا كتبنا هذه المعادلة كالتالي :

$$\Delta Y_t = \beta_1 + \beta_{2t} + u_t \qquad \text{(a.5.5.21)}$$

والتي تعني أن Y_t غير ساكنة



شكل (5.21) الاتجاه العام المحدد ضد الاتجاه العام العشوائي

المصدر: Charemze et al., op.cit., p.91

الاتجاه العام المحدد مع مركب ساكن AR(1) :

Deterministic Trend with Stationary AR(1) Component

إذا كان (1.5.21) $\beta_1 \neq 0$، $\beta_2 \neq 0$ و $\beta_3 < 1$ فإن لدينا

$$Y_t = \beta_1 + \beta_{2t} + \beta_3 Y_{t-1} + u_t \qquad \textbf{(6.5.21)}$$

والتي تعتبر ساكنة حول الاتجاه العام المحدد. لإدراك الفرق بين الاتجاه العام المحدد والاتجاه العام العشوائي، انظر شكل (5.21)[14]. السلسلة العشوائة في هذا الشكل مولدة بـ RWM: الطبيعي $Y_t = 0.5 + Y_{t-1} + u_t$ حيث إن 500 قيمة لـ u_t تم توليدها من التوزيع الطبيعي القياسي، والقيم المبدئية لـ Y تساوي 1. السلسلة المحددة تم توليدها كالتالي $Y_t = 0.5t + u_t$ حيث u_t تم توليده كما سبق، و t هو الزمن القياسي بترتيب زمني كما ترى من الشكل (5.21)، في حالة الاتجاه العام المحدد، الانحراف عن خط الاتجاه (والذي يمثل المتوسط غير الساكن) يعتبر تام العشوائية، وينتهي سريعًا، والذي يحدد بمكون الاتجاه العام $0.5t$. في حالة الاتجاه العام العشوائي، على العكس المكون العشوائي u_t يؤثر على قيم سلسلة Y على المدى البعيد.

(14) المناقشة التالية مأخوذة من

Wojciech W.charemaza et al., op. cit., pp. 89-91.

6.21 العمليات العشوائية المدمجة : INTEGRATED STOCHASTIC PROCESSES

نموذج السير العشوائي هو حالة خاصة من فئة أكثر عمومية من العمليات العشوائية، والمعروفة باسم العمليات المدمجة. تذكر أن RWM بدون اتجاه يعتبر غير ساكن، ولكن الفروق الأولى له كما هو موضح في (8.3.21) ساكنة. وبالتالي نسمى الـ RWM بدون اتجاه عملية مدمجة من الدرجة 1، ونرمز لها بالرمز $I(1)$. وبالمثل إذا كانت السلسلة الزمنية مأخوذة لها الفرق مرتين (بمعنى نأخذ الفرق الأولى للفروق الأولى) ساكنة سنسميها سلسلة زمنية مدمجة من الدرجة 2[15]. عمومًا، إذا كانت السلسلة الزمنية (غير الساكنة) لابد من أخذ الفروق لها d مرة لجعلها ساكنة، فإن هذه السلسلة الزمنية يقال عنها مدمجة من الدرجة d. السلسلة الزمنية Y_t المدمجة من الدرجة d يرمز لها بالرمز $Y_t \sim I(d)$. إذا كانت السلسلة الزمنية Y_t من الأصل ساكنة (بمعنى أنها لاتحتاج أخذ أي فروق لها) فإنها تسمى مدمجة من الدرجة صفر، ويرمز لها بالرمز $Y_t \sim I(0)$. وبالتالي سنستخدم المصطلح "سلسلة زمنية ساكنة" و "سلسلة زمنية مدمجة من الدرجة صفر" لنتحدث عن نفس الشئ.

معظم السلاسل الزمنية الاقتصادية هي في العموم $I(1)$: بمعنى أنها بوجه عام تصبح ساكنة بعد أخذ فروقها الأولى فقط. هل السلاسل الزمنية الموجودة في شكل (1.21 و2.21) $I(1)$ أو لها درجة أعلى؟ سنختبر ذلك في الفقرات 8.21 و 9.21.

خصائص السلسلة المدمجة : Properties of Integrated Series

الخصائص التالية للسلسلة الزمنية المدمجة:

دعنا نستعرض X_t، Y_t و Z_t ثلاث سلاسل زمنية.

1- إذا كانت $X_t \sim I(0)$ و $Y_t \sim I(1)$ فإن: $Z_t = (X_t + Y_t) = I(1)$، بمعنى أن أي توليفة خطية من سلسلة زمنية ساكنة وأخرى غير ساكنة تكون غير سالبة.

2 - إذا كان $X_t \sim I(d)$ فإن $Z_t = (a + bX_t) = I(d)$، بحيث إن a و ط ثوابت. بمعنى أن أي توليفة خطية من سلسلة $I(d)$ هي أيضًا $I(d)$، وبالتالي إذا كانت $X_t \sim I(0)$ فإن $Z_t = (a + bX_t) \sim I(0)$.

---

(15) على سبيل المثال، إذا كانت Y_t هي $I(2)$ إذن:

$$\Delta\Delta Y_t = \Delta(Y_t - Y_{t-1}) = \Delta Y_t - \Delta Y_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2}$$

والتي ستكون ساكنة. ولكن لاحظ أن: $\Delta\Delta Y_t = \Delta 2Y_t \neq Y_t - Y_{t-2}$.

3 – إذا كان $X_t \sim I(d_1)$ و $Y_t \sim I(d_2)$ فإن $Z_t = (aX_t + bY_t) \sim I(d_2)$ حيث $d_1 < d_2$.

4 – إذا كان $X_t \sim I(d)$ و $Y_t \sim I(d)$ فإن $Z_t = (aX_t + bY_t) \sim I(d^*)$ بحيث إن d^* عموماً مساوية لـ d ولكن في بعض الحالات تكون $d^* < d$ (انظر في الفقرة 11.21 الخاصة بالاندماج المزدوج.

كما ترى مما سبق، لابد من الحذر عند عمل توليفة بين اثنين أو أكثر من السلاسل الزمنية المدمجة عند درجات مختلفة.

لترى لماذا ذلك يعتبر أمراً ضرورياً، اعتبر نموذج الانحدار ذا المتغيرين، والذي ناقشناه في الفصل (3) وهو $Y_t = \beta_1 + \beta_2 X_t + u_t$. تحت صحة الفروض التقليدية للـ OLS فنحن نعلم:

$$\hat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2} \qquad (1.6.21)$$

حيث إن الحروف الصغيرة كالعادة، تشير إلى الانحرافات عن القيم المتوسطة، افترض أن Y_t هي $I(0)$ ولكن X_t هي $I(1)$، بالتالي فإن الأولى ساكنة، والأخيرة ليست ساكنة. وبما أن X_t ساكنة، فإن تباينها سيزداد، وبالتالي سيسود على البسط الموجود في المقدار (1.6.21) مما سينتج عنه أن $\hat{\beta}_2$ ستؤول إلى الصفر تقاربياً (بمعنى حدوث ذلك في العينات الكبيرة) ولن يكون لها حتى توزيع تقاربي[16].

7.21 ظاهرة الانحدار الزائف:
THE PHENOMENON OF SPURIOUS REGRESSION

لترى لماذا تأخذ السلسلة الزمنية الساكنة كل هذه الأهمية. اعتبر نموذجي السير العشوائيين التاليين:

$$Y_t = Y_{t-1} + u_t \qquad (1.7.21)$$

$$X_t = X_{t-1} + v_t \qquad (2.7.21)$$

حيث تم توليد 500 مفردة لـ u_t من $u_t \sim N(0, 1)$ و 500 مفردة لـ v_t من $v_t \sim N(0, 1)$ وافترضنا أن القيم المبدئية لكل من Y و X هي الصفر، وافترضنا أن u_t و v_t غير مرتبطين

(16) هذه النقطة مأخوذة من
MADDALLA ET AL., OP., CIT., P.26.

تسلسليًا، وغير مرتبطين تباعًا، وكما تعلم الآن فإن كلاً من هاتين السلسلتين غير ساكنتين، حيث إنهما $I(1)$ أو تظهر قيمهما اتجاه عام عشوائي.

افترض أننا نقوم بعمل انحدار لـ Y_t على X_t. بما أن Y_t و X_t غير مرتبطين عمليًا $I(1)$، فإن R^2 من أي انحدار لـ Y على X لابد أن تقترب من الصفر، بمعنى عدم وجود أي علاقة بين هذين المتغيرين. ولكن انتظر حتى ترى نتائج هذا الانحدار.

Variable	Coefficient	Std. error	t statistic
C	-13.2556	0.6203	-21.36856
X	0.3376	0.0443	7.61223
	R^2 = 0.1044	d = 0.0121	

كما ترى، فإن معامل الـ X له معنوية إحصائية عالية، على الرغم من انخفاض قيمة R^2 والتي لها معنوية إحصائية في عدم مساواتها بالصفر. من هذه النتائج، من الممكن أن تستنتج أن هناك علاقة إحصائية معنوية بين X و Y على الرغم من أن هذه الفكرة لم تكن موجودة من قبل. وهذا هو مضمون فكرة ظاهرة الانحدار الزائف. إن الانحدار الذي ليس له معنى والذي اكتشفه Yule(17).

أوضح Yule أن هذا الارتباط (الزائف) يظهر في السلاسل الزمنية غير الساكنة، حتى إذا كان حجم العينة كبيرًا جدًا. حيث يكون هناك شئ خاطئ في الانحدار، والذي يظهر في قيمة منخفضة لإحصاء Durbin- Watson d والذي يقترح وجود ارتباط ذاتي قوي من الدرجة الأولى، ووفقًا لـ Granger و Newblod فإن $R^2 > d$ يعتبر قاعدة جيدة لاكتشاف أن الانحدار المقدر يعتبر انحدارًا زائفًا، ويعتبر ماسبق مثالاً على ذلك.

وبالتالي، فنتائج الانحدار المعطاة أعلى تعتبر نتائج بدون معنى، وذلك يمكن رؤيته بسهولة من انحدار الفروق الأولى لـ Y_t ($=\Delta Y_t$) على الفروق الأولى لـ X_t ($=\Delta Y_t$) وتذكر أنه على الرغم من أن Y_t و X_t غير ساكنين، فإن فروقهما الأولى ساكنة. في مثل هذا الانحدار، ستجد أن R^2 عمليًا تساوي الصفر، كما يجب أن تكون، وقيم Durbin- Watson d حوالي 2. في تمرين 24.21 يسأل القارئ بأن يقوم بهذا الانحدار ويثبت العبارة سابقة الذكر.

(17) Yule, G. U., "Why Do We Sometimes Get Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Series," Journal of the Royal Statistical Society, vol. 89, 1926, pp. 1–64.

لمحاكاة واضحة Monte Carlo للانحدار الزائف انظر:

C.W.J. Granger and P. Newbold, "Spurous Regressions in Econometrics." Journal of Econometrics, vol. 2, 1976, pp. 111–120.

هذا المثال يعتبر تذكيراً قوياً للفرد، فإن يأخذ احتياطًا وحذراً شديداً عند القيام بتحليل انحدار لسلاسل زمنية تشتمل على اتجاهات عامة عشوائية. ولابد بالتالي أن يكون الفرد شديد الحذر عند قراءة نتائج انحدار يعتمد على متغيرات $I(1)$. وكمثال، انظر التمرين 26.21 وببعض القيود، فإن هذا صحيح بالنسبة للسلاسل الزمنية ذات الاتجاه العام المحدد، ومثال على ذلك في تمرين 25.21.

8.21 اختبارات السكون : TESTS OF STATIONARITY

حتى الآن محتمل أن يكون لدى القارئ فكرة جيدة عن طبيعة العمليات العشوائية الساكنة وأهميتها. في الواقع نحن نواجه سؤالين مهمين: (1) كيف نعرف ما إذا كانت سلسلة زمنية ما ساكنة أم لا؟ (2) إذا وجدناها غير ساكنة، هل هناك طريقة لجعلها ساكنة؟ سنجيب على السؤال الأول في الفقرة الحالية، ونناقش السؤال الثاني في الفقرة 10.21.

قبل البدء في ذلك، ضع في الاعتبار أننا مهتمون أساسًا بالسكون الضعيف أو سكون التغاير.

على الرغم من وجود العديد من اختبارات السكون، إلا أننا سنناقش منها فقط الاختبارات ذات الشهرة تاريخيًا. في هذه الفقرة، سنناقش اختبارين: (1) التحليل البياني و(2) اختبار مصور الارتباط. ونظرًا للأهمية الخاصة باختبار جذور الوحدة في الماضي القريب، سنناقشه في الفقرة التالية. سنشرح هذه الاختبارات بأمثلة مناسبة.

1- التحليل البياني : Graphical Analysis

كما لاحظنا من قبل، وقبل البدء في الاختبارات المختلفة، دائمًا ينصح الفرد برسم السلسلة الزمنية محل الدراسة، كما فعلنا في الشكلين (1.21 و 2.21) للبيانات المعطاة في جدول (1.21). مثل هذا الرسم، يعطي فكرة مبدئية عن الطبيعة المحتملة للسلسلة الزمنية، فمثلاً السلسلة الزمنية الخاصة بالـ GDP الموضحة في الشكل (1.21) سترى أنه بمرور الزمن، فإن GDP يزداد، حيث يتضح ذلك من الاتجاه العام المرتفع، مما قد يعني أن متوسط الـ GDP تغير.

مما قد يعني أن سلسلة GDP غير ساكنة. هذا قد يكون صحيحًا أو غير صحيح لباقي السلاسل الزمنية الاقتصادية للولايات المتحدة، الموضحة في الشكل (2.21). مثل هذا الحس المبدئي يعتبر نقطة البداية لاختبار أكثر دقة للسكون.

2- دالة الارتباط الذاتي (ACF) ومصور الارتباط:

2- Autocorrelation Function (ACF) and Correlogram

أحد الاختبارات البسيطة للسكون يعتمد على ما يسمى دالة الارتباط الذاتي (ACF) حيث إن الـ ACF عند الفترة الزمنية المتأخرة k والذي يرمز له بالرمز ρ_k معروف كالتالي:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$= \frac{\text{التغاير عند الفترة الزمنية المتأخرة } k}{\text{التباين}} \quad (1.8.21)$$

حيث إن التغاير عند الفترة الزمنية المتأخرة k، والتباين معرفين كما سبق. لاحظ العلاقة السابقة إذا كانت $k = 0$، $\rho_0 = 1$ (لماذا؟).

بما أن كلاً من التغاير والتباين مقاسان بنفس الوحدات القياسية، فإن ρ_k رقم صاف أو غير متأثر بوحدة القياس. حيث يقع بين 1- و 1 بالضبط كما هو الحال في معامل الارتباط. إذا رسمنا ρ_k ضد k، فإن الشكل الذي سنحصل عليه يسمى مصور ارتباط المجتمع.

بما أنه في الواقع العملي يكون لدينا تمثيل (عينة) من العملية العشوائية، فإننا نستطيع فقط حساب دالة الارتباط الذاتي من العينة (SAFC)، $\hat{\rho}_k$. لحساب ذلك، لابد أن نحسب أولاً تغاير العينة عند الفترة الزمنية المتأخرة k، $\hat{\gamma}_k$، وتباين العينة، $\hat{\gamma}_0$، والذي يعرف كالتالي[18]:

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n} \quad (2.8.21)$$

$$\hat{\gamma}_0 = \frac{\sum (Y_t - \bar{Y})^2}{n} \quad (3.8.21)$$

حيث n هو حجم العينة و$\bar{Y}$ متوسط العينة.

وبالتالي، فإن دالة الارتباط الذاتي في العينة عند الفترة الزمنية المتأخرة k هي:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (4.8.21)$$

(18) بشكل خاص لابد أن نقسم تغاير العينة عند الفترة الزمنية المتأخرة k على $(n-k)$ وتباين العينة على $(n-1)$ بدلاً من n (لماذا؟) حيث n هو حجم العينة.

والذي يمثل ببساطة النسبة بين تغاير العينة (عند الفترة الزمنية المتأخرة k) إلى تباين العينة. الشكل البياني الذي يتم فيه رسم $\hat{\rho}_k$ ضد k معروف باسم مصور ارتباط العينة. كيف يمكن لمصور ارتباط العينة أن يجعلنا نحدد ما إذا كانت سلسلة زمنية ما ساكنة أم لا؟ لهذا السبب دعنا أولاً نستعرض مصور ارتباط العينة لعملية عشوائية بحتة، ولعملية سير عشوائي. وبالعودة إلى RWM بدون اتجاه الموجود في (13.3.21). تم توليد عينة من 500 مقدار الخطأ، الـ u's، من التوزيع الطبيعي القياسي. مصور الارتباط للـ 500 خطأ العشوائي موضح في الشكل (6.21)، وقد أوضحنا مصور الارتباط حتى الـ 30 فترة زمنية متأخرة. وسنعلق لاحقًا على كيفية اختبار طول الفترة الزمنية المتأخرة. حاليًا انظر فقط إلى العمود المعنون AC، والذي يمثل دالة الارتباط الذاتي، والشكل الأول على اليسار والمعنون الارتباط الذاتي.

العينة: 2500

المفردات المتضمنة في الدراسة: 499

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	−0.022	−0.022	0.2335	0.629
		2	−0.019	−0.020	0.4247	0.809
		3	−0.009	−0.010	0.4640	0.927
		4	−0.031	−0.031	0.9372	0.919
		5	−0.070	−0.072	3.4186	0.636
		6	−0.008	−0.013	3.4493	0.751
		7	0.048	0.045	4.6411	0.704
		8	−0.069	−0.070	7.0385	0.532
		9	0.022	0.017	7.2956	0.606
		10	−0.004	−0.011	7.3059	0.696
		11	0.024	0.025	7.6102	0.748
		12	0.024	0.027	7.8993	0.793
		13	0.026	0.021	8.2502	0.827
		14	−0.047	−0.046	9.3726	0.806
		15	−0.037	−0.030	10.074	0.815
		16	−0.026	−0.031	10.429	0.843
		17	−0.029	−0.024	10.865	0.863
		18	−0.043	−0.050	11.807	0.857
		19	0.038	0.028	12.575	0.860
		20	0.099	0.093	17.739	0.605
		21	0.001	0.007	17.739	0.665
		22	0.065	0.060	19.923	0.588
		23	0.053	0.055	21.404	0.556
		24	−0.017	−0.004	21.553	0.606

		15	−0.037	−0.030	10.074	0.815
		16	−0.026	−0.031	10.429	0.843
		17	−0.029	−0.024	10.865	0.863
		18	−0.043	−0.050	11.807	0.857
		19	0.038	0.028	12.575	0.860
		20	0.099	0.093	17.739	0.605
		21	0.001	0.007	17.739	0.665
		22	0.065	0.060	19.923	0.588
		23	0.053	0.055	21.404	0.556
		24	−0.017	−0.004	21.553	0.606
		25	−0.024	−0.005	21.850	0.644
		26	−0.008	−0.008	21.885	0.695
		27	−0.036	−0.027	22.587	0.707
		28	0.053	0.072	24.068	0.678
		29	−0.004	−0.011	24.077	0.725
		30	−0.026	−0.025	24.445	0.752

شكل (6.21) مصور ارتباط الأخطاء العشوائية البحتة u، Ac ارتباط ذاتي، PAC= ارتباط ذاتي جزئي (انظر الفصل 22)، Q-Stat= إحصاء Q، prob= الاحتمال

الخط الرأسي المجسم في هذا الشكل، يمثل المحور الصفري، المشاهدات فوق هذا الخط تمثل قيمًا موجبة، والموجودة أسفل هذا الخط تمثل قيمًا سالبة. وكما هو واضح من هذا الشكل للعملية العشوائية البحتة، فإن الارتباط الذاتي عند فترات زمنية متأخرة عديدة يتأرجح حول الصفر. هذه هي صورة مصور الارتباط لسلسلة زمنية ساكنة، وبالتالي إذا كان مصور الارتباط لسلسلة زمنية حقيقية (اقتصادية) يشابه مصور الارتباط لسلسلة زمنية بحتة، فإننا نستطيع القول بأن هذه السلسلة الزمنية غالبًا ساكنة.

والآن دعنا نستعرض مصور ارتباط لسلسلة سير عشوائي، والذي تم توليده مثلاً كما في (13.3.21). الصورة الموضحة في شكل (7.21) تمثل ذلك. الصفة الأكثر وضوحًا من مصور الارتباط هذا، أن معاملات الارتباط الذاتي عند الفترات الزمنية المتأخرة المختلفة له قيم عالية جدًا حتى الفترة الزمنية المتأخرة الربع سنوية رقم 33. وفي الحقيقة، إذا اعتبرنا الفترات الزمنية المتأخرة الربع سنوية حتى الفترة رقم 66 فإن معاملات الارتباط الذاتي ستظل عالية أيضًا. فالمعاملات تكون حوالي 0.7 عند الفترة الزمنية المتأخرة 60 . (شكل 7.21) يعتبر مصور ارتباط مثاليًا لسلسلة زمنية غير ساكنة: معاملات الارتباط الذاتي تبدأ عند قيم مرتفعة جدًا، ويقل تدريجيًا ببطء ناحية الصفر، كلما زاد طول الفترة الزمنية المتأخرة.

العينة : 2500

المفردات المتضمنة : 499

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.992	0.992	493.86	0.000
		2	0.984	0.000	980.68	0.000
		3	0.976	0.030	1461.1	0.000
		4	0.969	0.005	1935.1	0.000
		5	0.961	−0.059	2402.0	0.000
		6	0.953	0.050	2862.7	0.000
		7	0.946	0.004	3317.3	0.000
		8	0.939	0.040	3766.4	0.000
		9	0.932	−0.009	4210.1	0.000
		10	0.927	0.055	4649.1	0.000
		11	0.921	0.018	5083.9	0.000
		12	0.916	0.039	5514.9	0.000
		13	0.912	0.002	5942.4	0.000
		14	0.908	0.056	6367.0	0.000
		15	0.905	0.061	6789.8	0.000
		16	0.902	0.000	7210.6	0.000
		17	0.899	0.006	7629.4	0.000
		18	0.896	0.030	8046.7	0.000
		19	0.894	0.053	8463.1	0.000
		20	0.892	0.013	8878.7	0.000
		21	0.890	−0.041	9292.6	0.000
		22	0.886	−0.040	9704.1	0.000
		23	0.882	−0.044	10113.	0.000
		24	0.878	−0.012	10518.	0.000
		25	0.873	−0.023	10920.	0.000
		26	0.867	−0.041	11317.	0.000
		27	0.860	−0.055	11709.	0.000
		28	0.853	−0.045	12095.	0.000
		29	0.846	−0.010	12476.	0.000
		30	0.839	0.008	12851.	0.000
		31	0.832	−0.006	13221.	0.000
		32	0.825	0.003	13586.	0.000
		33	0.819	−0.006	13946.	0.000

شكل (7.21) مصور ارتباط لسلسلة زمنية لسير عشوائي ، انظر الشكل (6.21) للتعريفات المختلفة للمتغيرات الموجودة في الشكل .

والآن دعنا نستعرض مثالاً اقتصاديًا تطبيقيًا. دعنا نختبر مصور ارتباط لسلسلة زمنية خاصة بالـ GDP المعطى في جدول (1.21). مصور الارتباط حتى الفترة الزمنية المتأخرة 25 يعطي شكلاً مماثلاً للمصور الارتباط الخاص بالسير العشوائي الموجود في شكل (7.21). فمعاملات الارتباط الذاتي تبدأ بقيم مرتفعة جدًا عند الفترة الزمنية المتأخرة الأولى (0.969)، وتنخفض تدريجيًا ببطء شديد. مما يعني أن السلسلة الزمنية الخاصة بالـ GDP تبدو غير ساكنة. إذا رسمنا مصور الارتباط لسلاسل زمنية اقتصادية أخرى للولايات المتحدة كما في الشكلين (1.21 و 2.21) سترى شكلاً عامًا مماثلاً مما يعني أن كل هذه السلاسل الزمنية غير مستقرة، وقد تكون غير مستقرة في القيمة المتوسطة أو في التباين أو في كليهما.

العينة : I-1970 إلى 4-1991

المفردات المتضمنة : 88

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.969	0.969	85.462	0.000
		2	0.935	–0.058	166.02	0.000
		3	0.901	–0.020	241.72	0.000
		4	0.866	–0.045	312.39	0.000
		5	0.830	–0.024	378.10	0.000
		6	0.791	–0.062	438.57	0.000
		7	0.752	–0.029	493.85	0.000
		8	0.713	–0.024	544.11	0.000
		9	0.675	0.009	589.77	0.000
		10	0.638	–0.010	631.12	0.000
		11	0.601	–0.020	668.33	0.000
		12	0.565	–0.012	701.65	0.000
		13	0.532	0.020	731.56	0.000
		14	0.500	–0.012	758.29	0.000
		15	0.468	–0.021	782.02	0.000
		16	0.437	–0.001	803.03	0.000
		17	0.405	–0.041	821.35	0.000
		18	0.375	–0.005	837.24	0.000
		19	0.344	–0.038	850.79	0.000
		20	0.313	–0.017	862.17	0.000
		21	0.279	–0.066	871.39	0.000
		22	0.246	–0.019	878.65	0.000
		23	0.214	–0.008	884.22	0.000
		24	0.182	–0.018	888.31	0.000
		25	0.153	0.017	891.25	0.000

شكل (8.21) مصور ارتباط للـ GDP الخاص بالولايات المتحدة I-1970 إلى IV-1991 .
انظر الشكل (6.21) للتعريفات المختلفة للمتغيرات الموجودة في الشكل .

سؤالان عمليان يجب الإجابة عليهما هنا. أولاً، كيف نختار طول الفترة الزمنية المتأخرة لحساب الـ ACF؟ ثانيًا، كيف يمكن معرفة ما إذا كان معامل الارتباط عند فترة زمنية متأخرة ما معنويًا إحصائيًا أم لا؟ الإجابة كالتالي.

اختيار طول الفترة الزمنية المتأخرة : The choice of lag length

هذا السؤال في الأساس سؤال تطبيقي. إحدى طرق الإجابة هو حساب الـ ACF من ثلث إلى ربع طول السلسلة الزمنية. وحيث إن لدينا بيانات اقتصادية عن 88 مفردة ربع سنويًا، فإن بهذه القاعدة تكون الفترات الزمنية المتأخرة مابين 22 و29 فترة زمنية ربع سنوية. النصيحة العملية الجيدة هي البدء بعدد فترات زمنية متأخرة كبير بشكل كاف ثم تقليلها بطريقة إحصائية ما مثل طريقة معلومات Akaike و Schwarz والتي سبق مناقشتها في الفصل (13). وكبديل يمكن استخدام الاختبارات الإحصائية التالية.

المعنوية الإحصائية لمعاملات الارتباط الذاتي :

Statistical Significance of Autocorrelation Coefficients

اعتبر على سبيل المثال مصور ارتباط للسلسلة الزمنية الخاصة بالـ GDP المعطى في الشكل (8.21). كيف يمكن تحديد ما إذا كان معامل الارتباط المساوي لـ 0.638 عند الفترة الزمنية المتأخرة 10 (ربع سنوية) معنوي إحصائيًا أم لا؟ المعنوية الإحصائية لأي $\hat{\rho}_k$ يمكن الحكم عليها من خلال خطئها القياسي. Bartlett أوضح أنه إذا كانت السلسلة الزمنية عشوائية بشكل تام أي لها عشوائية بحتة (انظر شكل 6.21) فإن معاملات الارتباط الذاتي للعينة $\hat{\rho}_k$ تكون تقريبًا[19].

$$\hat{\rho}_k \sim N(0, 1/n) \qquad (5.8.21)$$

بمعنى أنه في العينات كبيرة الحجم، فإن معاملات الارتباط الذاتي للعينة له التوزيع الطبيعي بمتوسط يساوي صفرًا وتباينًا يساوي واحدًا على حجم العينة. وبما أن لدينا 88 مفردة، فإن التباين يساوي 1/88=0.01136 والخطأ القياسي هو $\sqrt{0.01136}$ =01.066 . وبالتالي وفقًا لخصائص التوزيع الطبيعي القياسي، فإن الـ 95% فترة ثقة لأي ρ_k (المجتمع) هو:

$$\hat{\rho}_k \pm 1.96(0.1066) \qquad (6.8.21)$$

---

(19) M.S. Bartlett, "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series," Journal of the Royal Statistical Society, Series B, vol. 27, 1946, pp. 27–41.

وبشكل آخر ، فإن ذلك يعني أن :

$$\text{Prob}\,(\hat{\rho}_k - 0.2089 \leq \rho_k \leq \hat{\rho}_k + 0.2089) = 0.95 \qquad (7.8.21)$$

إذا احتوت فترة الثقة السابقة على الصفر ، فإننا لانستطيع رفض الفرض الخاص بأن القيمة الحقيقية لـ ρ_k هي الصفر، ولكن إذا لم تحتو هذه الفترة على 0، فإننا نرفض الفرض الخاص بأن القيمة الحقيقية لـ ρ_k هي الصفر. وبتطبيق ذلك على القيمة المقدرة $\hat{\rho}_{10}=0.638$ فإن القارئ يمكن أن يثبت أن 95% فترة الثقة للقيمة الحقيقية ρ_{10} هي (0.638 ± 0.2089) أو $(0.4291, 0.8469)$[20] وواضح أن فترة الثقة هذه لاتحتوي على الصفر، مما يعني أننا بدرجة ثقة 95%، فإن القيمة الحقيقية لـ ρ_{10} مختلفة إحصائيًا عن الصفر[21]. ويمكنك التأكد من أنه حتى عند الفترة الزمنية المتأخرة 20، فإن القيمة المقدرة لـ ρ_{20} لها معنوية إحصائية على المستوى 5%.

بدلاً من عمل اختبار معنوية لكل من معاملات الارتباط الذاتي على حدة، فإنه يمكن اختبار علم فرض مشترك على كل الـ ρ_k حتى فترة زمنية متأخرة ما ومساواة بالصفر آنيًا. يمكن القيام بذلك باستخدام إحصاء Q الذي اقترحه Box و Pierce والمعرف كالتالي[22].

$$Q = n\sum_{k=1}^{m}\hat{\rho}_k^2 \qquad (8.8.21)$$

حيث إن n = حجم العينة و m = طول الفترات الزمنية المتأخرة. الإحصاء Q يستخدم غالبًا كاختبار لما إذا كانت السلسلة الزمنية لها عشوائية بحتة أم لا في العينات كبيرة الحجم، هذا الإحصاء يتبع توزيع x^2 تقاربيًا بدرجات حرية m. عند التطبيق، إذا كانت القيمة المحسوبة لـ Q تزيد عن القيمة الحرجة لـ Q والمحسوبة من توزيع x^2 عند مستوى المعنوية المختار، فإن الفرد يرفض الفرض العدمي الخاص بأن

(20) حجم عينتنا المساوي لـ 88 مفردة على الرغم من أنه ليس حجم عينة كبيراً جدً،ا إلا أنه كبير بشكل يسمح باستخدام التقريب للتوزيع الطبيعي.

(21) وبشكل بديل، إذا قسمنا العينة المقدرة لأي ρ_k على الخطأ القياسي ($\sqrt{1/n}$) للعينات كبيرة الحجم نسبيًا، فإننا نحصل على القيمة القياسية Z والتي يمكن الحصول على احتمالها بسهولة من جدول التوزيع الطبيعي القياسي. وبالتالي للقيمة المقدرة $\rho_{10}=0.638$ فإن قيمة Z هي $0.638/0.1066=5.98$ (تقريبًا). إذا كانت القيمة الحقيقية فعلاً تساوي الصفر، فإن احتمال الحصول على قيمة Z مساوية لـ 5.98 أو أكبر صغير جداً وبالتالي نرفض الفرض الخاص بأن القيمة الحقيقية لـ ρ_{10} تساوي الصفر.

(22) G.E.P. Box and D.A. Pierce, "Distribution of Residual Autocorrelations in Autoregers-sive Integrated Moving Average Time Series Models," Journal of the American Statistical Association, vol. 65, 1970, pp. 1509–1526.

كل ρ_k (الحقيقية) تساوي الصفر، ويكون الفرض البديل في هذه الحالة هو على الأقل واحد من الـ ρ_k لايساوي الصفر. وإحصاء آخر مختلف عن إحصاء Q لـ Box-Pierce هو إحصاء Ljeng–Box (LB) والمعروف كالتالي[23].

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \sim \chi^2 m \qquad (9.8.21)$$

على الرغم من أنه في حجم العينة الكبيرة إحصاء Q و LB يتبعان توزيع x^2 بدرجات حرية m، إلا أن إحصاء LB وجد له خصائص أفضل في العينات صغيرة الحجم (أكثر قوة بالمعنى الإحصائي) من الإحصاء Q [24].

وبالعودة إلى مثال الـ GDP المعطى في شكل (8.21)، قيمة إحصاء الـ LB حتى الفترة الزمنية المتأخرة 250 حوالي 891.25. احتمال الحصول على قيمة مثل هذه القيمة لإحصاء LBs تحت صحة الفرض العدمي والخاص بكون مجموع مربعات معاملات الارتباط الذاتي المقدرة الـ 25 يساوي الصفر يساوي عمليًا الصفر، يتضح ذلك من العمود الأخير لهذه الأشكال السابقة. وبالتالي فنستنتج أن السلسلة الزمنية الخاصة بالـ GDP غير ساكنة، مما يدعم توقعنا من الشكل (1.21) بأن سلسلة الـ GDP قد تكون غير ساكنة. في تمرين 6.21 يسأل القارئ أن يثبت أن السلاسل الزمنية الأربع الاقتصادية الخاصة بالولايات المتحدة الأخرى هي أيضًا غير ساكنة.

9.21 اختبار جذر الوحدة : THE UNIT ROOT TEST

اختبار جذر الوحدة هو اختبار للسكون (أو عدم السكون)، والذي أصبح يستخدم بكثرة في السنوات العديدة الماضية. سنبدأ أولاً بشرح الاختبار، ثم توضيحه، مع استعراض بعض القيود الخاصة بهذا الاختبار.

نقطة البداية هي عملية جذر الوحدة (العشوائية) والتي ناقشناها في الفقرة 4.21 سنبدأ بالتالي.

$$Y_t = \rho Y_{t-1} + u_t \qquad -1 \le \rho \le 1 \qquad (1.4.21)$$

(23) G.M. Ljung and G.P.E. box, "On a Measure of Lack of Fit in Time Series Models," Biometrika, vol. 66, 1978, pp. 66–72.

(24) إحصاءات Q و LB قد لاتكون مناسبة في كل الحالات. للمزيد من التفاصيل عن ذلك انظر في: Maddala et. al., op. cit., p.19.

حيث u_t مقدار خطأ عشوائي بحت. نعمل أنه إذا كان $\rho=1$، بمعنى أننا في حالة جذر الوحدة، فإن (1.4.21) تصبح نموذج سير عشوائي بدون اتجاه، والذي عرفناه أنه عملية عشوائية غير ساكنة، وبالتالي لماذا لانقوم بعمل انحدار لـ Y_t على قيمتها المتأخرة زمنيًا Y_{t-1} (متأخرة فترة واحدة)، ونرى ما إذا كانت القيمة المقدرة لـ ρ تساوي إحصائيًا الـ 1 ؟ وإذا كانت كذلك، فإن Y_t تعتبر غير ساكنة. هذه هي فكرة عامة وراء اختبار جذر الوحدة للسكون.

لأسباب نظرية، قمنا بالتعامل مع (1.4.21) كالتالي: نطرح Y_{t-1} من كل من طرفي (1.4.21) لنحصل على:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + u_t \\ &= (\rho - 1)Y_{t-1} + u_t \end{aligned} \qquad (1.9.21)$$

والذي يمكن أن تكتب بشكل بديل كالتالي:

$$\Delta Y_t = \delta Y_{t-1} + u_t \qquad (2.9.21)$$

حيث $\delta = (\rho-1)$ و Δ كالعادة هي معامل الفرق الأولى.

في الواقع، من أجل ذلك، بدلاً من تقدير (1.4.21) فإننا نقدر (2.9.21) ونختبر الفرض (العدمي) القائل أن $\delta = 0$. إذا كانت $\delta = 0$ فإن $\rho=1$، وبالتالي فإن لدينا جذر الوحدة، مما يعني أن السلسلة الزمنية محل الاهتمام غير ساكنة. قبل البدء في تقدير (2.8.21) يمكن ملاحظة أنه إذا كانت $\delta = 0$ فإن (2.9.21) ستصبح كالتالي:

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \qquad (3.9.21)$$

وبما أن u_t هو مقدار خطأ عشوائي بحت، فهو ساكن، وذلك يعني أن الفروق الأولى لسلسلة زمنية لسير عشوائي ساكنة، وقد أوضحنا هذه النقطة من قبل.

الآن دعنا نعود إلى تقدير (2.9.21). وهذا أمر بسيط بشكل كاف، كل مانحتاج إلى فعله هو أخذ الفروق الأولى لـ Yt وعمل انحدار لها على Y_{t-1} ونرى ما إذا كان معامل الميل المقدر في هذا الانحدار ($\hat{\delta}$ =) يساوي الصفر أم لا. ما إذا كان يساوي الصفر، فإننا نستنتج أن Y_t غير ساكنة. ولكن إذا كانت قيمته سالبة فإننا نستنتج أن Y_t ساكنة[25]. السؤال الآن هو أي اختبار نستخدمه لمعرفة ما إذا كان معامل Y_{t-1} المقدر في (21.9.2) يساوي الصفر أم لا. يمكن أن تفكر في لماذا لاتستخدم اختبار t العادي؟

(25) Since $\delta = (\rho - 1)$, for stationarity must be less than one. for this to happen δ must be negative.

للأسف تحت صحة الفرض العدمي الخاص بأن $\delta = 0$ (أي أن $\rho = 1$)، قيمة t لمعامل Y_{t-1} المقدر لاتتبع توزيع t حتى في أحجام العينات الكبيرة، بمعنى ألا تتبع التوزيع الطبيعي تقاربيًا.

ما هو البديل إذن؟ Dickey و Fuller أوضحا أنه تحت صحة الفرض العدمي الخاص بـ $\delta = 0$، فإن قيمة t المقدرة لمعامل Y_{t-1} في (2.9.21) يتبع إحصاء τ (tau)(26). وقد قام هذان الكاتبان بحساب القيمة الحرجة لإحصاء tau على أساس محاكاة Monte Carlo. عينة من هذه القيم الحرجة معطاة في الملحق D، جدول (D.7). الجدول محدود ولكن Mackinnon قام بعمل جداول أكثر دقة ووضوح، والتي تستخدم الآن في العديد من الحزم الإلكترونية الاقتصادية(27). في التاريخ عرف إحصاء أو اختبار tau بأنه اختبار Dickey-Fuller (DF) تكريمًا لاكتشافهما لهذا الاختبار. والمثير للاهتمام هو أنه عند رفض الفرض الخاص بـ $\delta = 0$ (بمعنى أن السلسلة الزمنية ساكنة)، فإنه يمكننا استخدام اختبار t العادي (stuetend's).

الطريقة الفعلية لتطبيق اختبار DF تشتمل على العديد من القرارات. عندما ناقشنا طبيعة عملية جذر الوحدة في الفقرة 4.21 و 5.21، لاحظنا أن عملية السير العشوائي قد لايكون لها اتجاه، وقد يكون لها اتجاه أو يكون لها كل من اتجاه عام محدد وعشوائي. لتقدير اختبار DF في الحالات الممكنة المختلفة، فإن ذلك يتم بثلاثة أشكال مختلفة. بمعنى أنها تحت صحة فروض عدمية مختلفة.

(2.9.21) Y_t هي سير عشوائي: $\Delta Y_t = \delta Y_{t-1} + u_t$

(3.9.21) Y_t هي سير عشوائي باتجاه: $\Delta Y_t = (\beta_t + \delta Y_{t-1}) = u_t$

(4.9.21) Y_t هي سير عشوائي باتجاه ولها اتجاه عام عشوائي: $\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t$

حيث إن t هو الزمن أو متغير الاتجاه العام. في كل حالة يكون الفرض العدمي

---

(26) D. A. Dickey and W.A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association, vol. 74, 1979, pp 427–431.
انظر أيضًا: W. A. Fuller, Introduction to Statistical Time Series, John Wiley & Sons, New York, 1976.

(27) J. G. Mackinnon, "Critical Values of Cointegration Tests," in R. E. Engle and C. W. J. Granger, eds., Long-Run Econonic Relationships: Readings in Cointegration, Chap. 13, Oxford University Press, New ork, 1991.

هو أن $\delta = 0$، بمعنى أن جذر الوحدة موجود- السلسلة الزمنية غير ساكنة. الفرض البديل هو أن δ أقل من الصفر، بمعنى أن السلسلة الزمنية ساكنة[28]. إذا تم رفض الفرض العدمي، فإن ذلك يعني أن Y_t هي سلسلة زمنية ساكنة بمتوسط يساوي الصفر في الحالة (2.9.21) وتكون Y_t ساكنة بمتوسط لايساوي الصفر $[= \beta_1/(1-\rho)]$ في الحالة (3.9.21) وتكون Y_t ساكنة حول اتجاه عام محدد في (4.9.21).

من الضروري جداً أن نلاحظ أن القيم الحرجة لاختبار tau لاختبار الفرض القائل إن $\delta = 0$ مختلفة في كل من الحالات الثلاث السابقة لاختبار DF ويمكن رؤية ذلك بسهولة من ملحق D، جدول (D.7). والأكثر من ذلك، إذا كان على سبيل المثال، التوصيف (4.9.21) هو الصحيح، ولكننا قدرنا (2.9.21) سنكون قد وقعنا في خطأ التعريف أو التوصيف والذي تم استعراض عواقبه من قبل في الفصل (13). نفس الشيء قد يكون سليماً إذا قدرنا (3.9.21) بدلاً من الوضع الصحيح والموجود في (4.9.21). بالطبع لاتوجد طريقة لمعرفة أي التوصيفات سليم للبدء به. بعض أساليب المحاولة والخطأ ستكون ذات أهمية كبيرة فمعالجة البيانات والتعامل معها أمر لاحدود له.

عملية التقدير الفعلية تتم كالتالي: قدر (2.9.21) أو (3.9.21) أو (4.9.21) باستخدام الـ OLS، اقسم المعامل المقدر لـ Y_{t-1} في كل حالة على خطئه القياسي لحساب الإحصاء tau (τ)، وبالعودة إلى جداول DF (أو باستخدام أي حزم إلكترونية إحصائية)، إذا كانت القيمة المطلقة المحسوبة لإحصاء tau ($|\tau|$) تزيد عن القيمة الحرجة لـ DF أو Mac Kinnon فإننا نرفض الفرض العدمي القائل أن $\delta = 0$، وفي مثل هذه الحالة، تكون السلسلة الزمنية ساكنة. وعلى العكس إذا حسبنا (τ) وكانت لاتزيد عن القيمة الحرجة لـ tau، فإننا لانرفض الفرض العدمي، وفي مثل هذه الحالة تكون السلسلة الزمنية غير ساكنة. تأكد من أنك تستخدم القيمة الحرجة لـ τ المناسبة.

دعنا نرجع إلى السلسلة الزمنية الخاصة بالـ GDP الخاص بالولايات المتحدة. بالنسبة لهذه السلسلة، نتائج الانحدارات (2.9.21)، (3.9.21) و (4.9.21) هي كالتالي: المتغير التابع في كل حالة هو $\Delta Y_t = \Delta\ GDP_t$.

(28) استبعدنا إمكانية أن $\delta > 0$ حيث إن ذلك يعني أن $\rho > 1$ مما يعني أن السلسلة الزمنية محل الاهتمام ستتفجر.

$$\widehat{\Delta \text{GDP}}_t = 0.00576\text{GDP}_{t-1} \quad (6.9.21)$$
$$t = (5.7980) \qquad R^2 = -0.0152 \qquad d = 1.34$$

$$\widehat{\Delta \text{GDP}}_t = 28.2054 - 0.00136\text{GDP}_{t-1} \quad (7.9.21)$$
$$t = (1.1576) \quad (-0.2191) \qquad R^2 = 0.00056 \qquad d = 1.35$$

$$\widehat{\Delta \text{GDP}}_t = 190.3857 + 1.4776t - 0.0603\text{GDP}_{t-1} \quad (8.9.21)$$
$$t = (1.8389) \quad (1.6109) \quad (-1.6252)$$
$$R^2 = 0.0305 \qquad d = 1.31$$

اهتمامنا الأساسي هنا هو فيه t (=τ) لمعامل GDP_{t-1}. القيم الحرجة لـ τ عند 1 ، 5 و 10% للنموذج (6.9.21) هي 2.5897-، 1.9439- و 1.16177- بالترتيب. والقيم 3.5064-، 2.8947- و 2.5842- للنموذج (21.9.7) و 4.0661-، 3.4614- و 3.1567- للنموذج (8.3.21) كما لاحظنا من قبل، هذه القيم الحرجة مختلفة للنماذج الثلاثة.

قبل اختبار هذه النتائج، لابد أن نقرر أيًا من هذه النماذج الثلاثة يبدو مناسبًا. لابد من استبعاد النموذج (6.9.21) حيث إن معامل GDP_{t-1}، والذي يساوي δ يأخذ قيمة موجبة. ولكن حيث إن $\delta = (\rho-1)$، فإن قيمة موجبة لـ δ ستعني أن $\rho > 1$. وعلى الرغم من أن ذلك ممكن نظريًا، إلا أننا نستبعد هذه الحالة، حيث إن السلسلة الزمنية لـ GDP ستكون منفجرة[29] وبالتالي، فإنه متبقي الآن النموذجان (7.9.21) و(8.9.21). في كل من الحالتين القيمة المقدرة للمعامل δ تأخذ قيمة سالبة، مما يعني أن القيمة المقدرة لـ ρ أقل من 1. لهذين النموذجين، القيمة المقدرة لـ ρ هي 0.9986 و0.9397 بالترتيب. السؤال المهم الآن إذا كانت هذه القيم لها معنوية إحصائية أقل من 9، فإن ذلك يعني أن لدينا السلسلة الزمنية لـ GDP ساكنة.

بالنسبة للنموذج (7.9.21) القيم المقدرة لـ τ هي 0.2191- والتي تعتبر كقيمة مطلقة أقل حتى من القيمة الحرجة عند 10% والمساوية لـ 2.5842- . ونظرًا لأننا نستخدم القيم المطلقة، فإن الأولى أصغر من الأخيرة، وبالتالي نستنتج أن السلسلة الزمنية لـ GDP ليست ساكنة[30].

---

(29) المعنى الفني لذلك، حيث إن (2.9.21) هي معادلة فروق أولى، فإن الشرط المسمى شرط الاستقلال يتطلب أن $|\rho| < 1$.

(30) طريقة أخرى لصياغة ذلك هي أن قيمة τ المحسوبة، لابد أن تكون أكثر سالبية من قيمة τ الحرجة، وهذا ليس الحال هنا. وبالتالي نستنتج أن، طالما بوجه عام δ متوقع أن تكون سالبة، فإن إحصاء τ المقدر سيأخذ إشارة سالبة، وبالتالي قيمة سالبة كبيرة لـ τ عمومًا تعتبر مؤشرًا للسكون.

نفس الطريقة تنطبق على النموذج (8.9.21). قيمة τ المحسوبة هي 1.6252– وهي أقل من قيمة τ الحرجة عند 10% وهي 3.1567– في صورة قيم مطلقة. وبالتالي معتمدين على التحليل البياني، مصور الارتباط واختبار Dickey- Fuller فإن الاستنتاج النهائي هو أنه للفترات الربع سنوية من 1970 إلى 1991، فإن السلسلة الزمنية للـ GDP للولايات المتحدة غير ساكنة، بمعنى أنه تشتمل على جذر الوحدة.

اختبار Dickey- Fuller المزيد (ADF):

The Augmented Dickey- Fuller (ADF) test

عند القيام بعمل اختبار DF كما في (2.9.21)، (3.9.21) أو (4.9.21)، كان من المفترض أن مقدار الخطأ u_t غير مرتبط. ولكن في الحالات التي يكون فيها u_t مرتبطًا قام Fuller و Dickey بعمل اختبار جديد معروف باسم اختبار Dickey- Fuller المزيد. هذا الاختبار يتم "بزيادة" المعادلات الثلاث السابقة، بإضافة قيم في فترات زمنية متأخرة للمتغير التابع ΔY_1. بالتحديد، افترض أننا استخدمنا (4.9.21). فإن اختبار ADF يتم عن طريق تقدير الانحدار التالي:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \qquad \textbf{(9.9.21)}$$

حيث إن ε_t هو مقدار الخطأ العشوائي الصافي البحت و $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$، $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$ وهكذا. عدد مقادير فروق الفترات الزمنية المتأخرة والذي يجب احتواؤه في هذا الانحدار يتم اختباره عادة عمليًا، الفكرة وراء تضمين مقادير كافية هي جعل مقدار الخطأ في (9.9.21) غير مرتبط تسلسليًا. في ADF سنظل نختبر ما إذا كان $\delta=0$ أم لا، واختبار ADF يتبع نفس التوزيع التقاربي لإحصاء DF، وبالتالي يتم استخدام نفس القيم الحرجة.

لإعطاء لمحة عن هذه الطريقة. قدرنا (9.9.21) لسلسلة GDP باستخدام فترة زمنية متأخرة واحدة للفروق الخاصة بـ GDP. النتائج كالتالي[31].

$$\begin{aligned} \widehat{\Delta \text{GDP}}_t &= 234.9729 + 1.8921t - 0.0786\text{GDP}_{t-1} + 0.3557\Delta\text{GDP}_{t-1} \\ t &= \quad (2.3833) \qquad (2.1522) \qquad (-2.2152) \qquad\qquad (3.4647) \qquad \textbf{(10.9.21)} \\ & R^2 = 0.1526 \qquad d = 2.0858 \end{aligned}$$

(31) فروق الفترات الزمنية المتأخرة من درجة أعلى تم دراستها ولكن كانت غير معنوية.

قيمة t ($=\tau$) لمعامل GDP_{t-1} ($=\delta$) هي 2.2152– ولكن هذه القيمة كقيمة مطلقة أقل بكثير من حتى القيمة الحرجة لـ t عند 10% والتي تساوي 3.1570–، ومرة أخرى، فإن ذلك يعني أنه بعد التعامل مع الارتباط الذاتي المحتمل في مقدار الخطأ، فإن سلسلة الـ DGP تظل غير ساكنة.

اختبار معنوية أكثر من معامل واحد: (اختبار F):

Testing The Significance of more than one Coefficient: The F test

افترض أننا قدرنا النموذج (5.9.21)، ونريد اختبار الفرض القائل بأن $\beta_1=\beta_2=0$ بمعنى أن النموذج هو RWM بدون اتجاه وبدون اتجاه عام أيضًا. لاختبار هذا الفرض المشترك، يمكن أن نستخدم اختبار F المقيد الذي ناقشناه في الفصل (8). أي أننا نقدر (5.9.21) (الانحدار غير المقيد) ونقدر (5.9.21) بعد حذف الجزء المقطوع من المحور الصادي والاتجاه العام. ثم نستخدم اختبار F المقيد كما هو موضح في المعادلة (9.7.8) باستثناء أننا لانستطيع استخدام جدول الـ F التقليدي للحصول على قيم F الحرجة كما فعلوا في إحصاء t، فإن Dickey و Fuller عملا قيم F الحرجة للتعامل مع مثل هذه المواقف، عينة من تلك القيم معطاة في الملحق D، جدول (D.7)، وفي تمرين 27.21 يوجد مثال على ذلك.

اختبارات جذر الوحدة (PP)[32]: The Phillips- Perron (PP) unit root tests

فرض مهم في اختبار DF خاص بمقادير الأخطاء u_t، حيث يفترض أنها مستقلة وموزعة بشكل متماثل. اختبار ADF يعتبر تعديلاً لاختبار DF حتى يمكن التعامل مع حالة الارتباط المتسلسل المحتمل في مقدار الخطأ عن طريق إضافة فروق الفترات الزمنية المتأخرة إلى المتغير المنحدر عليه. Phillips و Perron استخدما طرقًا إحصائية لامعلمية ليتعاملا مع مشكلة الارتباط المتسلسل في مقادير الأخطاء بدون إضافة مقادير الفروق في الفترات الزمنية المتأخرة. وبما أن التوزيع التقاربي لاختبار PP هو نفس التوزيع الخاص بإحصاء ADF، فإننا لن نتعمق أكثر من ذلك في هذه النقطة.

---

(32) P.C. Phillps and P. Perron, "Testing for a Uit Root in Time Series Regression," Biometrka, vol. 75, 1988, pp. 335–346.

اختبار PP متوافر الآن في عديد من برامج الحزم الإلكترونية.

نقد اختبارات جذر الوحدة[33]: A Critique of the unit root tests

ناقشنا العديد من اختبارات جذر الوحدة، ومازال هناك العديد منها. السؤال الآن هو: لماذا يوجد العديد من اختبارات جذر الوحدة؟ الإجابة تنبع من حجم وقوة مثل هذه الاختبارات. فحجم الاختبار يعني مستوى المعنوية (أي احتمال حدوث الخطأ من النوع الأول)، وقوة الاختبار تعني احتمال رفض الفرض العدمي وهو خاطئ. قوة الاختبار تحسب عن طريق طرح احتمال الخطأ من النوع الثاني من 1، الخطأ من النوع الثاني هو احتمال قبول فرض عدمي خاطئ. أكبر قيمة للقوة هي 1. معظم اختبارات جذر الوحدة تعتمد على أن الفرض العدمي الخاص بالسلسلة الزمنية محل الاهتمام يفترض أن لها جذراً يساوي الوحدة، وبالتالي فالسلسلة تكون غير ساكنة، الفرض البديل هو أن السلسلة الزمنية ساكنة.

حجم الاختبار: دعنا نتذكر من الفصل (13) المناقشة التي قمنا بها حول مستوى المعنوية الحقيقي والاسمي. اختبار DF يعتبر اختباراً حساساً للطريقة التي نفذ بها. تذكر أننا ناقشنا ثلاث حالات مختلفة لاختبار DF: (1) سير عشوائي خالص، (2) سير عشوائي باتجاه، (3) سير عشوائي باتجاه وله اتجاه عام أيضاً. إذا كان على سبيل المثال النموذج الصحيح هو (1) ولكننا قدرنا (2) واستنتجنا مثلاً أنه عند مستوى معنوية 5% السلسلة الزمنية ساكنة، هذا الاستنتاج قد يكون خطأ، حيث إن مستوى المعنوية الحقيقي في هذه الحالة أكبر بكثير من 5%[34]. تحريف حجم الاختبار قد يحدث أيضاً نتيجة استبعاد مكون متوسط متحرك (MA) من النموذج (انظر الفصل 22 للمزيد عن المتوسط المتحرك).

قوة الاختبار: معظم اختبارات من نوع الـ DF لها قوة ضعيفة، بمعنى أنها تميل إلى قبول فرض جذر الوحدة بشكل أكثر تكراراً مما يجب. بمعنى أن هذه الاختبارات قد تجد جذر الوحدة حتى ولو كان فعلاً غير موجود. هناك العديد من الأسباب لذلك:

---

(33) لمزيد من التفاصيل، انظر: TEREENCE C. MILLS, OP. CIT, PP. 87-88

(34) التجربة Monte Carlo خاصة بذلك، انظر: Charemza et al., Op. Cit., P. 114.

أولاً: قوة الاختبار تعتمد على مدى (زمن) البيانات أكثر من مجرد حجم العينة. فبمعلومية حجم العينة n، فإن قود الاختبار تكون أعلى عندما يزداد المدى، وبالتالي فإن اختبار أو اختبارات جذر الوحدة المعتمدة على 30 مفردة على مدار 30 عامًا قد يكون لها قوة أكبر من نظيرها المعتمد مثلاً على 100 مفردة على مدار 100 يوم.

ثانيًا: إذا كانت $\rho \approx 1$ ولكن ليست مساوية بالضبط لـ 1، فإن اختبار جذر الوحدة قد يثبت أن مثل هذه السلسلة الزمنية غير ساكنة.

ثالثًا: هذه الأنواع من الاختبارات تفترض وجود جذر وحدة وحيد، بمعنى أنهم يفترضون أن السلسلة الزمنية هي I (1) ولكن إذا كانت السلسلة الزمنية مدمجة عند درجة أعلى من 1، مثلاً، I (2) فإنه سيكون هناك أكثر من جذر وحدة واحد. وفي هذه الحالة الأخيرة، من الممكن استخدام اختبار Dickey- Pantula.[35]

رابعًا: إذا كانت هناك انكسارات بنائية في السلسلة الزمنية (انظر الفصل الخاص بالمتغيرات الوهمية) بسبب مثلاً إفطار على زيت OPEC، فإن اختبارات جذر الوحدة قد لاتظهر ذلك.

عند تطبيق اختبارات جذر الوحدة، لابد أن يضع الفرد في الاعتبار الحدود والقيود الموجودة على استخدام مثل هذه الاختبارات. بالطبع تم عمل بعض التعديلات على مثل هذه الاختبارات، مثل التي قام بها Perron و Ng, Elliot, Rothenberg و Leybaunce و Stock, Fuller[36]. بسبب ذلك أيد كل من Maddala وKim عدم استخدام اختبارات DF، ADF، PP التقليدية.

وذلك وراء الحدوث الآن، حيث إن معظم حزم البرامج الإلكترونية الخاصة بالاقتصاد القياسي تحتوي على اختبارات جديدة. ولكن يجب القول بأنه حتى الآن لايوجد اختبار منتظم أكبر قوة لاختبارات جذر الوحدة.

---

(35) D.a. Dickey and S. Pantula, "Determining the Order of Differencing in Autoregressive Processers, "Journal of Business and Economic Statistics, vol. 5, 1987, pp. 455–461.

(36) لمعرفة المزيد عن هذه الاختبارات، انظر Maddala et al., op. cit., Chap. 4.

10.21 تحويل السلاسل الزمنية غير الساكنة :
TRANSFORMING NONSTATIONARY TIME SERIES

والآن، بعد أن استعرضنا المشاكل المرتبطة بالسلاسل الزمنية غير الساكنة، فالسؤال العملي الآن هو: ماذا نفعل في مثل هذا الموقف. لتجنب مشكلة الانحدار الزائف والتي قد تنشأ من انحدار سلسلة زمنية غير ساكنة على واحدة أو أكثر عن السلاسل الزمنية غير الساكنة، لابد أن نقوم بعمل تحويل للسلسلة الزمنية غير الساكنة لجعلها ساكنة. طرق التحويل تعتمد على ما إذا كانت السلسلة الزمنية ساكنة للفروق (DSP) أو ساكنة في الاتجاه العام (TSP) أم لا. سنتناول كلاً من هذه الطرق على التوالي.

العمليات الساكنة ذات الفروق : Difference Stationary processes

إذا كانت السلسلة الزمنية لها جذر الوحدة، فإن الفروق الأولى لهذه السلسلة ساكنة[37]. وبالتالي الحل هنا هو أخذ الفروق الأولى لهذه السلسلة الزمنية، بالعودة إلى السلسلة الزمنية الخاصة بالـ GDP في الولايات المتحدة، قد رأينا بالفعل أن هذه السلسلة لها جذر الوحدة. دعنا نرى الآن ماذا سيحدث عن أخذ الفروق الأولى لسلسلة الـ GDP.

دع $\Delta GDP_t = (GDP_t - GDP_{t-1})$. للتسهيل دع $D_t = \Delta GDP_t$. والآن اعتبر الانحدار التالي:

$$\begin{aligned} \widehat{\Delta D_t} &= 16.0049 \quad - \quad 0.06827D_{t-1} \\ t &= (3.6402) \quad (-6.6303) \\ R^2 &= 0.3435 \qquad d = 2.0344 \end{aligned} \tag{1.10.21}$$

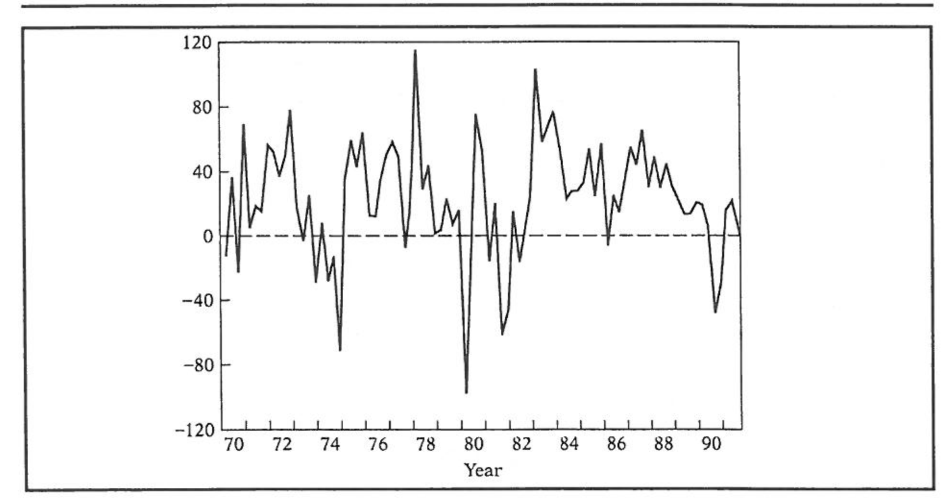
القيمة الحرجة عند 1% لقيمة DFτ هي 3.5073-. وبما أن قيمة τ المحسوبة ($=t$) أكثر سالبية من القيمة الحرجة، فإننا نستنتج أن الفروق الأولى لـ GDP ساكنة، بمعنى أنها $I(0)$. كما هو موضح في شكل (9.21). إذا قارنت شكل (9.21) مع شكل (10.21)، سترى الفرق الواضح بين الاثنين.

العملية الساكنة ذات الاتجاه العام : Trend- Stationary process

كما رأينا في الشكل (5.21)، فإن TSP تكون ساكنة حول خط الاتجاه العام، وبالتالي الطريقة الأبسط لجعل هذه السلسلة الزمنية ساكنة، هي عمل انحدار لها على الزمن، وبواقي هذا الانحدار ستكون ساكنة.

---

(37) إذا كانت السلسلة الزمنية هي $I(2)$، ستحتوي على جذرين وحدة. في مثل هذه الحالة، سنحتاج إلى أخذ الفروق مرتين. إذا كانت $I(d)$ لابد من أخذ الفروق d مرة حيث إن d عدد صحيح.



شكل (9.21) الفروق الأولى لـ GDP الخاص بالولايات المتحدة (ربع سنوية) 1970–1991

بشكل آخر، قم بعمل الانحدار التالي:

$$Y_t = \beta + \beta_2 t + u_t \qquad (2.10.21)$$

حيث إن Y_t هي السلسلة الزمنية محل الدراسة، و t هي متغير الاتجاه العام مقاس بشكل فيه ترتيب زمني.

الآن: (3.10.21)

$$\hat{u}_t = (Y_t - \hat{\beta}_1 - \hat{\beta}_2 t)$$

ستكون ساكنة. $\hat{u}_t$ معروف باسم سلسلة زمنية متعلقة بالاتجاه العام (الخطية)، من المهم ملاحظة أن الاتجاه العام قد يكون غير خطي. على سبيل المثال، قد يكون:

$$Y_t = \beta_1 + \beta_2 t + \beta_3 t^2 + u_t \qquad (4.10.21)$$

والتي تعتبر سلسلة اتجاه عام تربيعية. إذا كانت هذه هي الحالة، فإن البواقي من (4.10.21) ستكون الآن سلسلة زمنية ذات اتجاه عام (تربيعي).

لابد أن نشير إلى أنه، إذا كانت السلسلة الزمنية DSP ولكن عاملناها على إنها TSP، فإن ذلك يسمى عمل فروق بأقل مما يجب. على الجانب الآخر، إذا كانت السلسلة الزمنية TSP ولكن عاملناها على إنها DSP فإن ذلك يسمى عمل فروق بأكثر مما يجب. التوابع الخاصة بمثل هذه الأخطاء في التعريف، قد تكون خطيرة، وذلك يعتمد على كيفية التعامل مع خواص الارتباط التسلسلي الموجودة في مقادير الأخطاء الناتجة[38].

(38) لمزيد من التفاصيل، انظر Maddala et al., op. cit., sec. 2.7

يمكن أن نلاحظ بشكل عابر أن معظم السلاسل الزمنية الخاصة بالاقتصاد القياسي الكلي هي DSP أكثر من أن تكون TSP.

11.21 الاندماج المزدوج.. انحدار سلسلة زمنية ذات جذر الوحدة على سلسلة زمنية أخرى لها جذر الوحدة أيضًا: COINTEGRATION: REGRESSION OF A UNIT ROOT TIME SERIES ON ANOTHER UNIT ROOT TIME SERIES

قد حذرنا من قبل، بأن انحدار سلسلة زمنية غير ساكنة على سلسلة زمنية أخرى غير ساكنة، قد يؤدي إلى وجود انحدار زائف. دعنا الآن نفترض أن لدينا PCE و PDI سلسلتان زمنيتان موجودتان في جدول (1.21). وبعمل تحليل جذر الوحدة لكل من هاتين السلسلتين على حدة، ستجد أن كليهما $I(1)$، أي أنهما يحتويان على جذر الوحدة. افترض إذن أننا قمنا بعمل انحدار لـ PCE على PDI كالتالي:

$$\mathrm{PCE}_t = \beta_1 + \beta_2 \mathrm{PDI}_t + u_t \qquad (1.11.21)$$

دعنا نكتب ذلك كالتالي:

$$u_t = \mathrm{PCE}_t - \beta_1 - \beta_2 \mathrm{PDI}_t \qquad (2.11.21)$$

دعنا نقم الآن بتحليل جذر الوحدة لـ u_t ووجدنا أنه ساكن. أي أنه $I(0)$، فهذا يكون موقفًا مثيرًا للاهتمام، فعلى الرغم من أن PCE_t و PDI_t كل منهما على حدة $I(1)$ أي أن لهما اتجاه عشوائي، فإن توليفتيهما الخطية (2.11.21) هي $I(0)$. أي أن التوليفة الخطية ألغت الاتجاه العام العشوائي في كل من السلسلتين. إذا اعتبرت الاستهلاك والدخل كمتغيرين $I(1)$، الادخار معرف على أنه (الدخل- الاستهلاك) قد يكون $I(0)$.

وكنتيجة لذلك انحدار الاستهلاك على الدخل كما في (1.11.21) يكون له معنى (أي غير زائف). في مثل هذه الحالة، يقال إن هذين المتغيرين بينهما اندماج مزدوج. وبعبارات اقتصادية، فإن أي متغيرين يقال إن لهما اندماج مزدوج إذا كانت بينهما علاقة توازنية أو علاقة ما على المدى البعيد. النظرية الاقتصادية التي استخدمت مصطلحات التوازن مثل نظرية الكمية لـ Fisher أو نظرية تساوي الشراء (PPP) وهكذا.

باختصار باعتبار أننا تأكدنا من بواقي الانحدارات مثل (1.11.21) وكانت I(0) أو ساكنة، فإن طرق الانحدار التقليدية (بما فيها من اختبارات t و F) والتي درسناها من قبل، تكون مناسبة للتطبيق على بيانات السلاسل الزمنية (غير الساكنة). الإسهام القيم لمصطلحات جذر الوحدة، الاندماج المزدوج إلى ما غير ذلك، أجبرنا على تحديد ما إذا كانت بواقي الانحدار ساكنة أم لا. وكما لاحظ Granger "اختبار الاندماج المزدوج يمكن النظر له على أنه اختبار أولي لتجنب موقف (الانحدار الزائف)"[39].

وباستخدام **مصطلحات نظرية الاندماج المزدوج**، فإن انحداراً مثل الموجود في (1.11.21) يعرف باسم انحدار الاندماج المزدوج، ومعامل الميل β_2 يعرف باسم معلمة الاندماج المزدوج. مفهوم الاندماج المزدوج يمكن أن يطبق أيضًا على نماذج الانحدار التي تحتوي على k متغير منحدر. وفي مثل هذه الحالة، سيكون لدينا k معلمة للاندماج المزدوج.

اختبار الاندماج المزدوج: Testing for Cointegration

عدد من الطرق التي تستخدم في اختبار الاندماج المزدوج تم استعراضها من قبل تاريخيًا سنعتبر هنا طريقتان بسيطتان متنافستان وهما: (1) اختبار جذر الوحدة الـ DF أو ADF على البواقي المقدرة من انحدار الاندماج المزدوج و (2) اختبار Durbin-Watson لانحدار الاندماج المزدوج[40] (CRDW).

اختبار(EG) Engle-Granger أو الاختبار المزيد لـ Engle-Granger (AEG)

Engle- Granger (EG) or Augmented Engle- Granger (AEG) Test

نعرف بالفعل كيف نطبق اختبارات جذر الوحدة لـ DF أو ADF. كل مانحتاج فعله هو تقدير انحدار مثل (1.11.21)، نحصل على البواقي، ثم نستخدم اختبارات

---

(39) C.W.J. Granger, "Developments in the Study of Co-Integrated Economic Variables," Oxford Bulletin of Economics and Statistics, vol. 48, 1986, p. 226.

(40) هناك فرق بين اختبارات جذر الوحدة واختبارات الاندماج المزدوج. كما لاحظ david A. Dickey، dennis W. Jansen و daniel I. Thornton "اختبارات جذر الوحدة تم عملها على السلاسل الزمنية ذات المتغيرات، حيث كل منها له جذر الوحدة (غير مشروط)". انظر مقالتهم:
"A Primer on Cointegration with an Application to Money and Income", Economic Review, Federal Reserve Bank of St. Louis, March-April 1991, P.59
كما يقترح الاسم، فإن هذه المقالة مقدمة ممتازة لاختبارات الاندماج المزدوج.

DF أو ADF[41]. هناك شيء ما لابد من الاحتراس منه، فبما أن القيمة المقدرة لـ u_t معتمدة على معامل الاندماج المزدوج المقدر β_2، فإن القيم الحرجة المعنوية لـ DF ولـ ADF تعتبر غير مناسبة. Engle و Granger حسبا هذه القيم وهي موجودة في المراجع[42]. وبالتالي فاختبارات DF و ADF في الكتاب الحالي معروفة باسم اختبارات Engle- Granger (EG) أو اختبارات Engle- Granger (AEG) المزيدة. عمومًا الآن يوجد العديد من حزم البرامج الإلكترونية الموجودة بها هذه القيم الحرجة مع بعض النتائج الأخرى.

دعنا نستعرض هذه الاختبارات. سنقوم أولاً بعمل انحدار لـ PCE على PDI ونحصل على الانحدار التالي:

$$\widehat{\text{PCE}}_t = -171.4412 + 0.9672\text{PDI}_t$$
$$t = (-7.4808) \quad (119.8712) \qquad \textbf{(3.11.21)}$$
$$R^2 = 0.9940 \qquad d = 0.5316$$

بما أن PCE و PDI كل منهما غير مستقر بشكل منفرد، فهناك إمكانية أن يكون هذا الانحراف زائفًا. لكن عندما قمنا بعمل اختبار جذر الوحدة على البواقي التي حصلنا عليها من (3.11.21)، حصلنا على النتائج التالية:

$$\widehat{\Delta \hat{u}_t} = -0.2753\hat{u}_{t-1}$$
$$t = (-3.7791) \qquad \textbf{(4.11.21)}$$
$$R^2 = 0.1422 \qquad d = 2.2775$$

القيمة الحرجة لـ τ عند 1% لـ Engle-Granger هي 2.5899–. وبما أن قيمة τ $(=t)$ المحسوبة أكثر سالبية من هذه القيم، فإننا نستنتج أن بواقي انحدار PCE على PDI هي I (0)، بمعنى أنه ساكن. وبالتالي (3.11.21) يمثل انحدار اندماج مزدوج. وهذا الانحدار ليس زائفًا على الرغم من أن هذين المتغيرين كل منهما على حدة غير ساكن. وبالتالي يمكن تسمية (3.11.21) باسم دالة الاستهلاك على المدى البعيد أو الساكن، ويتم تفسير معالمهما كمعالم في المدى البعيد. وبالتالي فإن 0.9672 تمثل الميل الحدي المتوازن أو الميل الحدي في المدى البعيد للمستهلك (MPC).

---

(41) إذا لم يوجد اندماج مزدوج بين PEC و PDI فأي توليفة خطية منهما ستكون غير ساكنة، وبالتالي تكون u_t غير ساكنة أيضًا.

(42) R.F. Engle and C.W. Granger, "Co-integration and Error Correction: Representation, Estimation and Testing," Econometrica, vol. 55, 1987, pp. 251- 276.

اختبار Durbin- Watson لانحدار الاندماج المزدوج :

Cointegrating Regression Durbin- Watson (CRDW) test

اختبار CRDW هو طريقة بديلة أسرع في معرفة ما إذا كان PCE و PDI بينهما اندماج مزدوج أم لا، ويعتبر Sargon و Bhargave أول من قدما القيم الحرجة الخاصة بهذا الاختبار(43). في CRDW تستخدم Durbin-Watson d التي نحصل عليها من انحدار الاندماج المزدوج، مثل $d=$ 0.5316 المعطاة في (3.11.21) ولكن الآن يكون الفرض العدمي هو $d = 0$ بدلاً من الفرض التقليدي $d = 2$. ويرجع ذلك إلى أنه في الفصل (12) وجد أن $d \approx 2\ (1-\hat{\rho})$، وبالتالي إذا كان هناك جذر الوحدة، فإن قيمة ρ المقدرة ستكون تقريبًا 1، مما يعني أن d ستكون تقريبًا صفرًا.

على أساس 10.000 محاكاة مكونة من 100 مفردة، فإن القيم الحرجة 1، 5 و10% لاختبار الفرض الخاص بأن القيمة الحقيقية لـ $d = 0$ هي 0.511، 0.386 و 0.322 بالترتيب. وبالتالي إذا كانت قيمة d المحسوبة أقل مثلاً من 0.511 فإننا نرفض الفرض العدمي للاندماج المزدوج عند مستوى 1%. في مثالنا الحالي، قيمة 0.5316 أكبر من هذه القيمة الحرجة مما يعني أن PCE و PDI بينهما اندماج مزدوج مما يتطابق مع ما توصلنا إليه من قبل عن استخدام اختبار EG(44).

لتجميع كل استنتاجاتنا ووفقًا لكل من اختبارات CRDW و EG، فإن PCE و PDI بينهما اندماج مزدوج(45). فعلى الرغم من أن كلاً منهما على حدة يمثل سيراً عشوائيًا إلا أن بينهما علاقة مستقرة في المدى البعيد، فلن يبتعدا عن بعضهما البعض، وذلك واضح في الشكل (1.21).

---

(43) J.D. Sargan and A. S. Bhargava, "Testing Residuals from Least-Squares Regression for being Generaed by the Gaussian Random Walk," Econometrica, vol. 51, 1983, pp. 153-174.

(44) هناك بعض الشكوك حول تفوق CRDW على DF، وموجود ذلك بالتفصيل في المراجع. هذا الشك يكمن حول قوة هذين الاختبارين أي احتمال الوقوع في الخطأ من النوع الثاني. Engle و Granger على سبيل المثال يفضلان اختبار ADF عن اختبار CRDW

(45) اختبارات EG و CRDW يوجد الآن بدائل عنها باختبارات أكثر قوة مثل تلك التي قام بها Johansen. ولكن استعراض طريقة Johansen تقع خارج نطاق هذا الكتاب. حيث إن الرياضيات التي تشملها هذه الطريقة معقدة نوعا ما، وعلى الرغم من ذلك، يوجد الآن العديد من حزم البرامج الإلكترونية التي تستخدم طريقة Johansen.

الاندماج المزدوج وأسلوب تصحيح الخطأ (ECM):
Cointegration and Error Correction Mechanism (ECM)

أثبتنا أن PCE و PDI بينهما اندماج مزدوج، بمعنى أنه توجد علاقة توازنية بين الاثنين على المدى البعيد. بالطبع في المدى القصير قد يوجد عدم توازن. وبالتالي يمكن التعامل مع مقدار الخطأ في (2.11.21) كأنه "خطأ التوازن". ويمكن أن تستخدم مقدار الخطأ هذا لربط السلوك في المدى القصير لـ PCE مع قيمته في المدى البعيد. لربط أسلوب تصحيح الخطأ (ECM) استخدم أولاً عن طريق Sargon[46] ثم شهره بعد ذلك Engle و Granger عندما استخدماه في تصحيح التوازن. نظرية مهمة معروفة باسم نظرية التمثيل لـ Granger والتي تنص على أنه إذا كان المتغيران X و Y بينهما اندماج مزدوج، فإن العلاقة بين الاثنين يمكن التعبير عنها بـ ECM. لفهم ما تعنيه هذه العبارة، دعنا نسترجع مثالنا الخاص بـ PCI- PDI. والآن اعتبر النموذج التالي:

$$\Delta \mathrm{PCE}_t = \alpha_0 + \alpha_1 \Delta \mathrm{PDI}_t + \alpha_2 u_{t-1} + \varepsilon_t \qquad (5.11.21)$$

حيث Δ كالعادة ترمز إلى معامل الفروق الأولى، ε_t مقدار الخطأ العشوائي $u_t = (\mathrm{PCE}_{t-1} - \beta_1 - \beta_2 \mathrm{PDI}_{t-1})$، أي فترة زمنية واحدة متأخرة للخطأ من انحدار الاندماج المزدوج الموجود في (1.11.21).

معادلة ECM (5.11.21) تعني أن ΔPCE تعتمد على ΔPDI، وأيضًا على مقدار خطأ التوازن[47]. إذا كان الأخير لا يساوي الصفر، فإن النموذج لا يوجد فيه توازن. افترض أن ΔPDI يساوي الصفر و u_{t-1} موجب. فإن ذلك يعني أن PCE_{t-1} له قيمة عالية جدًا، بحيث لا يمكن أن يكون في وضع توازن. بمعنى أن PCE_{t-1} أعلى من قيمته التوازنية المساوية لـ $(\alpha_0 + \alpha_1 \mathrm{PDI}_{t-1})$ وبما أن α_2 متوقع أن تكون سالبة،

(46) J.D. Sargon, "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in K. J. Wallis and D. F. Hendry, eds., Quantitative Economics and Econometric Analysis, Basil Blackwell, Oxford, U.K., 1984.

(47) المناقشة التالية تعتمد على:
Gary Koop, op. cit., pp. 159- 160 and Kerry Peterson, op. cit., Sec. 8.5.

فالمقدار $\alpha_2 u_{t-1}$ سيكون سالبًا ، وبالتالي ΔPCE_t سيكون سالبًا لاسترجاع التوازن. بمعنى أنه إذا كان PCEt أعلى من قيمته التوازنية فإنه سيبدأ في النزول في الفترة التالية لتصحيح خطأ التوازن كما في التسمية ECM. وبنفس الطريقة إذا كان u_{t-1} سالبًا (بمعنى أن PCE أقل من قيمته التوازنية) فإن $\alpha_2 u_{t-1}$ سيكون موجبًا مما سيجعل ΔCPE_t موجبًا ويرفع قيمة PCE_t في الفترة t.

وبالتالي، القيمة المطلقة لـ α_2 تحدد متى سيتم استرجاع وضع التوازن. في الواقع نقدر u_{t-1} عن طريق $\hat{u}_{t-1} = (PCE_t - \hat{\beta}_1 - \hat{\beta}_2 PDI_t)$.

بالعودة إلى مثالنا التوضيحي، فإن الوضع التطبيقي لـ (5.11.21) هو:

$$\begin{aligned} \widehat{\Delta PCE}_t &= 11.6918 + 0.2906\Delta PDI_t - 0.0867\hat{u}_{t-1} \\ t &= (5.3249) \quad (4.1717) \quad (-1.6003) \\ R^2 &= 0.1717 \quad d = 1.9233 \end{aligned} \qquad \textbf{(6.11.21)}$$

إحصائيًا، مقدار خطأ التوازن يساوي الصفر، مما يعني أن PCE تتعدل وفقًا لتغيرات PDI في نفس الفترة الزمنية. وكما توضح (6.11.21) فإن التغيرات قصيرة المدى في PDI لها تأثير موجب على تغيرات المدى القصير في الاستهلاك الشخصي. ويمكن أن يتم تفسير 0.2906 كميل حدي للاستهلاك في المدى القصير (MPC)، MPC في المدى البعيد معطى عن طريق الإحصاء المقدر لعلاقة التوازن (3.11.21) ويساوي 0.9672 .

وقبل الوصول إلى الاستنتاج العام من هذه الفقرة، فإن التحذير الذي قاله S. G. Hall يجب وضعه في الاعتبار وهو:

في حين أن مفهوم الاندماج المزدوج له أهمية نظرية في تدعيم نموذج تصحيح الخطأ، إلا أنه مازال هناك عدد من المشاكل المحيطة بالتطبيق العملي، القيم الحرجة وأداء هذه الاختبارات عندما يكون حجم العينة صغيرًا غير معروف بالنسبة لعدد كبير من هذه النماذج، مع العلم بأن استخدام مصور الارتباط قد يظل أداة مهمة في هذا الموضوع[48].

(48) S. G. Hall, "An Application of the Granger and Engle Two-Step Estimation Procedure to the United Kingdom Aggregate Wage Data," Oxford Bulletin of Economics and Statistics, vol. 48, no. 3, August 1986. p. 238. See also John Y. Campbell and Pierre Perron, "Pitfalls and Oppo

12.21 بعض التطبيقات الاقتصادية : SOME ECONOMIC APPLICATIONS

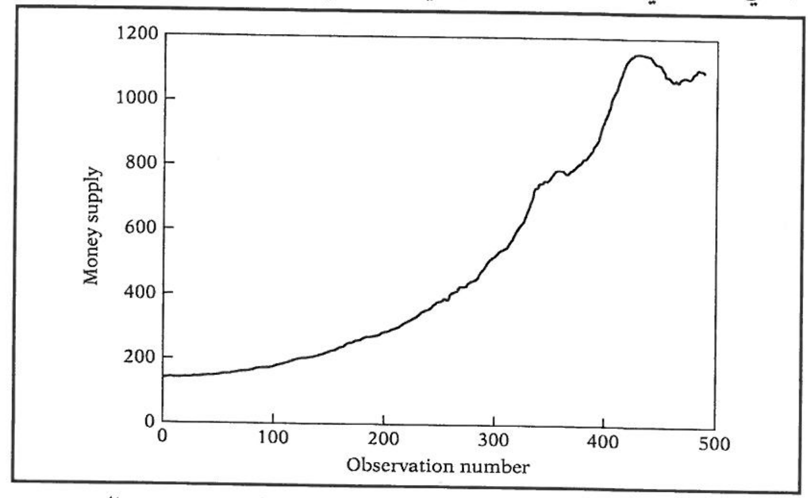
سنختم هذا الفصل ببعض الأمثلة العملية .

مثال 1.21

المعروض شهريًا من المال M1 في الولايات المتحدة ، يناير 1951 إلى 30 سبتمبر 1999

M. Monthly money supply in the United States, January 1951 to September 30, 1999

شكل (10.21) يوضح المعروض من المال M1 للولايات المتحدة في الفترة من يناير 1951 إلى سبتمبر 1999 . من معرفتنا بالسكون ، يتضح أن السلسلة الزمنية للمعروض من المال M1 غير ساكنة ، والذي يمكن إثباته باستخدام تحليل جذر الوحدة (لاحظ أنه : لضيق المساحة لم نستعرض البيانات الأصلية ، والتي يمكن الحصول عليها من لجنة الاحتياطي الفيدرالي أو من بنك الاحتياطي الفيدرالي لـ St. Louis) .



شكل (10.21) المعروض من المال في الولايات المتحدة خلال الفترة 01 : 1951 إلى 09 : 1999

$$\Delta\hat{M}_t = 0.2618 + 0.0159t - 0.0044M_{t-1}$$

$$t = (0.7919) \quad (4.4227) \quad (-3.0046) \qquad (1.12.21)$$

$$R^2 = 0.0670 \qquad d = 0.7172$$

القيم الحرجة لـ τ والخاصة بالنسب 1 ، 5 ، 10% هي 3.9811− ، 3.4210− و 3.1329− بالترتيب . وبما أن قيمة t هي 3.0046− أقل سالبية من هذه القيم الحرجة t، فإننا نستنتج أن السلسلة الزمنية الخاصة بـ M1 غير ساكنة ، بمعنى أن هذه السلسلة الزمنية تحتوي على جذر الوحدة ، أو إنها سلسلة زمنية من نوع I(1) . حتى عندما يوجد العديد من قيم المتغير في فترات زمنية متأخرة لـ ΔM_t (من ADF) فالاستنتاج لايتغير . على الجانب

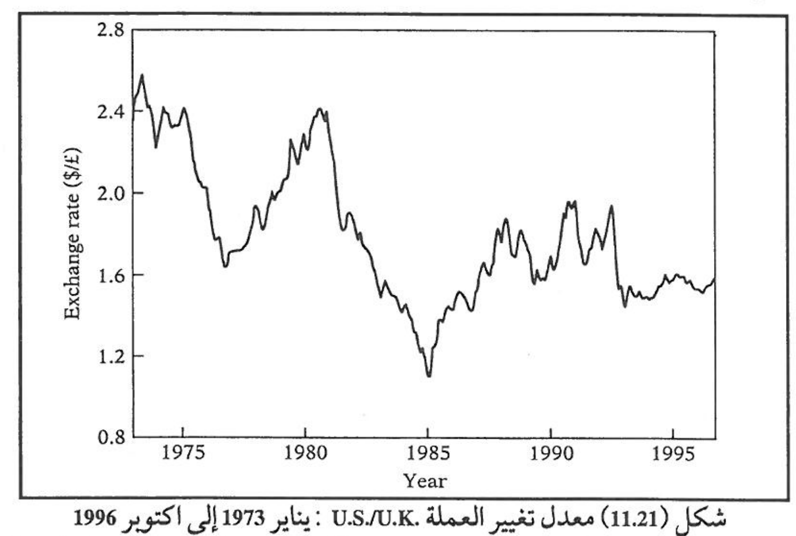
الآخر، الفروق الأولى للمعروض من المال M1 تعتبر سلسلة زمنية ساكنة (تأكد من صحة ذلك).

مثال 2.21 :

معدل تغيير العملة بين الولايات المتحدة/ المملكة المتحدة : 1 يناير 1973 إلى 10 أكتوبر 1996

The U.S./ U.K. Exchange Rate: January, 1973 to October 10, 1996

شكل (11.21) يوضح معدل تغيير العملة ($/£) من يناير 1973 إلى أكتوبر 1996 بجملة مشاهدات 286 مشاهدة. الآن يمكنك معرفة أن هذه السلسلة الزمنية تعتبر سلسلة زمنية غير ساكنة. وبإجراء اختبار جذر الوحدة، نحصل على القيم التالية لإحصاء J: ‎-1.2749 (بدون جزء مقطوع من المحور الصادي، بدون اتجاه عام)، ‎-1.7710 (يوجد جزء ثابت مقطوع من المحور الصادي)، ‎-1.6269 (يوجد اتجاه عام). كل هذه القيم كقيم مطلقة أقل من القيم الحرجة لـ τ والتي نحصل عليها من جداول DF المناسبة، وذلك يؤكد الانطباع الذي يتكون لدى القارئ من خلال الرسم البياني والخاص بأن معدل تغيير العملة بين .U.S./U.K يعتبر سلسلة زمنية غير ساكنة.



شكل (11.21) معدل تغيير العملة .U.S./U.K : يناير 1973 إلى اكتوبر 1996

مثال 3.21 :

مؤشر سعر المستهلك في الولايات المتحدة (CPI)، يناير 1947 إلى يناير 2000

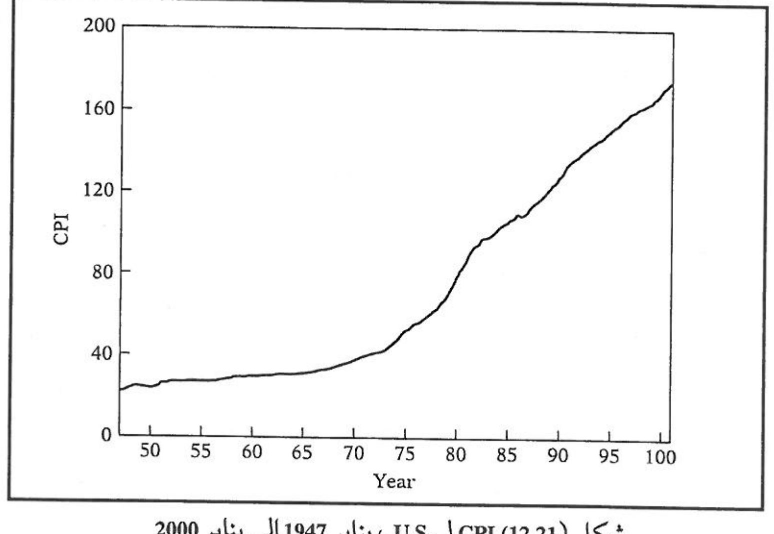
U.S. Consumer price index (CPI), January 1947 to January 2000

شكل (12.21) يوضح الـ CPI للولايات المتحدة في الفترة من يناير 1947 إلى يناير 2000 لعدد 649 مفردة. السلسلة الزمنية الخاصة بـ CPI مثل سلسلة M1 التي تم دراستها

سابقًا، توضح اتجاهًا عامًا متزايدًا لأعلى. اختبار جذر الوحدة أعطى النتائج التالية:

$$\widehat{\Delta CPI_t} = -0.0094 + 0.00051t - 0.00066CPI_{t-1} + 0.5473\Delta CPI_{t-1} \quad (2.12.21)$$
$$t = (-0.6538) \quad (4.3431) \quad (-1.5472) \quad (16.4448)$$
$$R^2 = 0.5177 \quad d = 2.1410$$

قيمة t(=τ) لـ CPI_{t-1} هي 1.5472−. القيمة الحرجة بنسبة 10% هي 3.1317−. ووفقًا للقيم المطلقة فإن τ المحسوبة أقل من قيمتها الحرجة τ وبالتالي نستنتج أن CPI ليست سلسلة زمنية ساكنة. ويمكن أن توصف بأن لها اتجاهًا عامًا متغيراً (لماذا؟). عمومًا إذا أخذنا الفروق الأولى لسلسلة CPI، ستجد أنها ساكنة وبالتالي فإن CPI هو سلسلة زمنية ساكنة للفروق (DC).



شكل (12.21) CPI لـ U.S.، يناير 1947 إلى يناير 2000

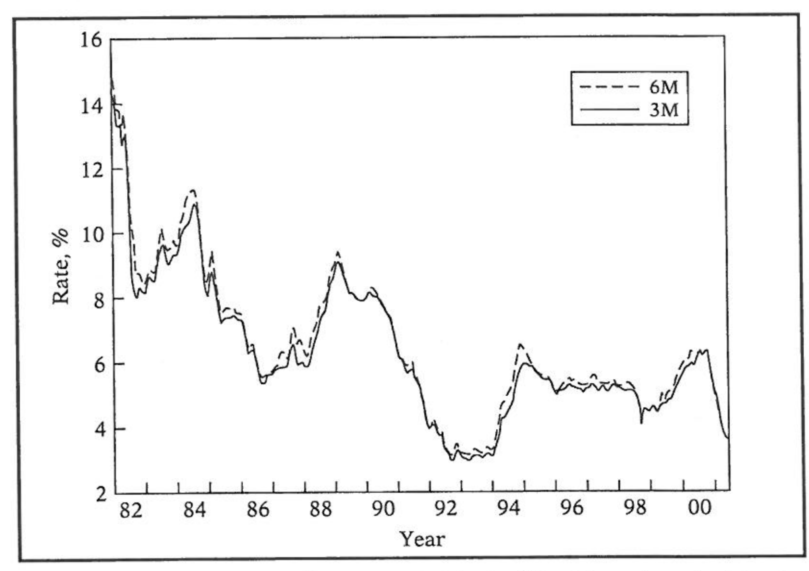
مثال 4.21 :

هل معدل الورقة المالية لمدة 3 أشهر ولمدة 6 أشهر بينهما اندماج مزدوج؟

Are 3 month and 6 month treasury bill rates cointegrated?

شكل (13.21) يوضح معدلات الأوراق المالية (T bill) للولايات المتحدة ذات الـ 3 أشهر، وأخرى ذات 6 أشهر في الفترة من يناير 1982 إلى يونيو 2001 لمجموع 234 مفردة. هل الشكل يوضح أن هذين المعدلين بينهما اندماج مزدوج: أي أنه توجد علاقة توازنية بين الاثنين؟

من النظرية المالية، فإننا نتوقع حدوث ذلك، وإلا فإن المحكمين سيستغلون أي تضارب بين معدلات المدى القريب ومعدلات المدى البعيد. أولاً، دعنا نرى ما إذا كانت هاتان السلسلتان ساكنتين.



شكل (13.21) معدلات الأوراق المالية لثلاثة وستة أشهر (استحقاق دين ثابت)

وفقًا لنموذج السير العشوائي الخالص (أي بدون جزء ثابت مقطوع من المحور الصادي وبدون اتجاه عام) فإن كلاً من المعدلين ساكنان. بإضافة الجزء المقطوع من المحور الصادي والاتجاه العام وفرق وحدة واحدة متأخرة زمنيًا، فإن النتائج تجعلنا نستنتج أن هذين المعدلين لهما اتجاه عام ساكن، حيث إن معامل الاتجاه العام في كلتا الحالتين سيكون سالبًا ومعنويًا عند حوالي 7%. وبالتالي وفقًا لأي من النتائج ستستخدم فإن المعدلين إما ساكنان أو لهما اتجاه عام ساكن.

عند عمل انحدار لمعدل T bill ذات 6 أشهر على المعدل ذي الـ 3 أشهر فإننا نحصل على الانحدار التالي:

$$\widehat{TB6}_t = -0.0456 + 1.0466TB3_t$$
$$t = (-1.1207) \quad (171.6239) \qquad R^2 = 0.9921 \qquad d = 0.4055 \qquad (3.12.21)$$

بتطبيق اختبار جذر الوحدة، فإن بواقي الانحدار السابق ستكون ساكنة، مما يعني أن معدل الأوراق المالية ذات الـ 3 أشهر والـ 6 أشهر بينهما اندماج مزدوج.

باستخدام هذه المعلومة، حصلنا على نموذج تصحيح الخطأ التالي (ECM):

$$\Delta\widehat{TB6}_t = -0.0067 + 0.9360\Delta TB3_t - 0.2030\hat{u}_{t-1}$$
$$t = (-0.8662) \quad (41.9592) \qquad (-5.3837) \qquad (4.12.21)$$
$$R^2 = 0.8852 \qquad d = 1.5604$$

حيث إن $\hat{u}_{t-1}$ هو مقدار تصحيح الخطأ في فترة زمنية متأخرة عن الفترة السابقة. كما توضح هذه النتائج فإن 0.2 من التعارض بين المعدلين في الشهر السابق تم حذفه هذا الشهر(49). بالإضافة لذلك، فإن التغيرات قصيرة المدى في معدلات الأوراق المالية

ذات الـ 3 أشهر تكون أكثر انعكاساً في معدلات الأوراق المالية ذات الـ 6 أشهر ، حيث إن معامل الميل بين الاثنين هو 0.936 . وهذه النتيجة تعتبر متوقعة بسبب كفاءة سوق المال في الولايات المتحدة.

13.21 الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

1 - يفترض ضمنيًا تحليل الانحدار المعتمد على بيانات سلاسل زمنية، أن السلسلة الزمنية محل الاهتمام تعتبر ساكنة. اختبارات t و F التقليدية تعتمد على هذا الافتراض.

2 - في الواقع معظم السلاسل الزمنية الاقتصادية غير ساكنة.

3 - العملية العشوائية يقال عنها إنها ضعيفة السكون إذا كان متوسطها، تباينها وتغايرها الذاتي ثابتين مع مرور الزمن (بمعنى أنهما غير متغيرين زمنيًا).

4 - بشكل غير أساسي، فإن السكون الضعيف يمكن اختباره بمصور الارتباط للسلسلة الزمنية، وهو عبارة عن شكل بياني للارتباط الذاتي عند فترات زمنية متأخرة عديدة. السلسلة الزمنية الساكنة يكون فيها مصور الارتباط متناقصًا تدريجيًا بشكل سريع، في حين أنه بالنسبة للسلسلة الزمنية غير الساكنة فإنه يزول أن ينخفض بشكل أكثر بطئًا. بالنسبة للسلسلة تامة العشوائية، فإن الارتباط الذاتي عند كل الفترات الزمنية المتأخرة بوحدة واحدة أو أكبر يكون مساويًا للصفر.

5 - بشكل أساسي، فإن السكون يمكن اختباره عن طريق معرفة ما إذا كانت السلسلة الزمنية لديها جذر وحدة أم لا. اختبارات Dickey- Fuller (DF) واختبارات Dickey- Fuller المزيدة (ADF) يمكن أن تستخدم لهذا الغرض.

6 - السلسلة الزمنية الاقتصادية يمكن أن تكون ساكنة ولها اتجاه عام (TS)، أو ساكنة الفروق (DS). السلسلة الزمنية TS يكون لها اتجاه عام محدد، أما السلسلة الزمنية DS فإن لها اتجاهًا عامًا عشوائيًا أو متغيرًا. التطبيق المشترك من إدخال الزمن أو متغير الاتجاه العام في نموذج الانحدار لإضافة اتجاه العام للبيانات يتم لتعديل فقط

(49) بما أن كلاً من معدلات الأوراق المالية هي في الفترة الحالية، فإن ذلك يعني أن معدل TB في 6 أشهر كان أعلى من معدل TB في 3 أشهر أكثر مما كان متوقعًا في الشهر الماضي، وهذا الشهر ستقل بنسبة 0.2 للاحتفاظ بالعلاقة في المدى البعيد بين معدلي الفائدة. لفهم المزيد عن النظرية الأساسية للعلاقة بين معدلات الفائدة في المدى القريب والمدى البعيد انظر في أي كتاب خاص بالبنوك والمال تحت أي فقرة أو فصل خاص بمعدلات الفائدة.

السلاسل الزمنية من نوع TS. اختبارات DF و ADF يمكن تطبيقها لمعرفة ما إذا كانت السلسلة الزمنية هي من نوع TS أو من نوع DS.

7 - انحدار متغير سلسلة زمنية ما على متغير سلسلة زمنية واحد أو أكثر قد يؤدي إلى نتائج غير منطقية أو زائفة. وهذه الظاهرة معروفة باسم الانحدار الزائف. أحد طرق التغلب على هذه المشكلة هو معرفة ما إذا كان هناك اندماج مزدوج بين السلاسل الزمنية أم لا.

8 - الاندماج المزدوج يعني أنه على الرغم من أن السلاسل الزمنية منفردة غير ساكنة، فإن التوليفة الخطية من اثنين أو أكثر من هذه السلاسل الزمنية يمكن أن يكون ساكنًا. اختبارات EG، AEG و CRDW يمكن استخدامها لمعرفة ما إذا كانت سلسلتان زمنيتان أو أكثر بينهما اندماج مزدوج أم لا.

9 - الاندماج المزدوج بين أي سلسلتين زمنيتين (أو أكثر) يعني أن هناك علاقة توازنية على المدى البعيد بينهما.

10 - طريقة تصحيح الخطأ (ECM) والتي قام بها Engle و Granger تعني إعادة توفيق سلوك متغير اقتصادي ما في المدى القريب مع سلوكه في المدى البعيد.

11 - مجال السلاسل الزمنية في الاقتصاد القياسي يتطور تدريجيًا. النتائج المعطاة والاختبارات في بعض الحالات تكون مؤقتة وتحتاج لمزيد من العمل. السؤال المهم والذي يحتاج إلى إجابة هو: لماذا تكون بعض السلاسل الزمنية الاقتصادية ساكنة وبعضها الأخرى تكون غير ساكنة.

تمارين : EXERCISES

أسئلة : Questions

1.21 ما هو المقصود بالسكون الضعيف؟

2.21 ما المقصود بالسلسلة الزمنية المدمجة؟

3.21 ما معنى جذر الوحدة؟

4.21 إذا كانت السلسلة الزمنية I(3). كم عدد الفروق اللازم أخذها لجعل السلسلة الزمنية ساكنة.

5.21 ماهو اختبار Dickey- Fuller (DF) واختبار DF المزيد؟

6.21 ماهو اختبار Engle- Granger (EG) واختبار EG المزيد؟

7.21 ماهو المقصود بالاندماج المزدوج؟

8.21 ماهو الفرق، إذا كان هناك فرق، بين اختبارات جذر الوحدة واختبارات الاندماج المزدوج؟

9.21 ماهو الانحدار الزائف؟

10.21 ماهي العلاقة بين الاندماج المزدوج والانحدار الزائف؟

11.21 ماهو الفرق بين الاتجاه العام المحدد والاتجاه العام العشوائي؟

12.31 ماهو المقصود بالعملية ساكنة الاتجاه العام (TSP) والعملية ساكنة الفروق (DSP)؟

13.21 ما هو نموذج السير العشوائي؟

14.21 " بالنسبة لعملية سير عشوائي، التباين غير محدود" . هل توافق على هذه العبارة؟ لماذا؟

15.21 ما هي طريقة تصحيح الخطأ (ECM)؟ وماهي علاقتها مع الاندماج المزدوج؟

مسائل : Problems

16.21 باستخدام البيانات الموجودة في جدول (1.21)، احصل على مصور الارتباط حتى الفترة المتأخرة رقم 25 للسلسلة الزمنية لكل من PCE، PDI، الأرباح، والمقسم على المساهمين. ماهو الشكل العام الذي تراه؟ أي من هذه السلاسل الزمنية يبدو أنه ساكن؟

17.21 لكل سلسلة زمنية في تمرين 16.21، استخدم اختبار DF لمعرفة ما إذا كانت هذه السلسلة الزمنية تحتوي على جذر الوحدة أم لا. إذا كان هناك جذر الوحدة، كيف يمكنك وصف هذه السلسلة الزمنية؟

18.21 استكمالاً على تمرين 17.21. كيف يمكنك أن تقرر ما إذا كان اختبار ADF مناسبًا أكثر للاستخدام عن اختبار DF؟

19.21 اعتبر السلسلة الزمنية الخاصة بالأرباح، والأرباح المقسمة على المساهمين المعطاة في جدول (1.21). بما أن الأرباح المقسمة على المساهمين تعتمد على الأرباح. اعتبر النموذج البسيط التالي:

$$\text{Dividends}_t = \beta_1 + \beta_2 \text{Profits} + u_t$$

(a) هل تعتقد أن هذا الانحدار سيعاني من ظاهرة الانحدار الزائف؟ لماذا؟

(b) هل هناك اندماج مزدوج بين الأرباح والأرباح المقسمة على المساهمين؟ كيف يمكنك اختبار ذلك صراحة؟ إذا كان، بعد إجراء الاختبار، هناك اندماج مزدوج، هل ستتغير إجابتك على السؤال a؟

(c) طبق طريقة تصحيح الخطأ (ECM) لدراسة سلوك الأرباح المقسمة في المدى القريب والبعيد وعلاقتها مع الأرباح.

(d) إذا اختبرت كلاً من الأرباح والأرباح المقسمة كلاً على حدة، هل تجد اتجاهًا عامًا عشوائيًا أم محددًا؟ ماهي الاختبارات التي ستستخدمها؟

(e)(*) افترض أن هناك اندماجًا مزدوجًا بين الأرباح والأرباح المقسمة. وبالتالي بدلاً من عمل انحدار للأرباح المقسمة على الأرباح، فقد قمت بعمل انحدار للأرباح على الأرباح المقسمة. هل هذا الانحدار الأخير يمكن القيام به؟ أي هل يعتبر انحدارًا سليمًا؟

20.21 احصل على الفروق الأولى للسلاسل الزمنية المعطاة في جدول (1.21) وارسمها. احصل أيضًا على مصور الارتباط لكل سلسلة زمنية حتى الفترة الزمنية المتأخرة رقم 25. ما الذي تلاحظه من مصورات الارتباط؟

21.21 بدلاً من عمل انحدار للأرباح المقسمة على الأرباح، افترض أنك قمت بعمل انحدار للفروق الأولى للأرباح المقسمة على الفروق الأولى للأرباح. هل سيوجد جزء ثابت مقطوع من المحور الصادي في هذا الانحدار؟ علل إجابتك. وضح الخطوات الحسابية اللازمة.

(*) اختياري.

22.21 استكمالاً للتمرين السابق. كيف يمكنك اختبار سكون انحدار الفروق الأولى؟ في المثال الحالي، ما الذي تتوقعه كوضع سابق ولماذا؟ وضح كل الخطوات الحسابية اللازمة.

23.21 من قطاع السكن الخاص بالمملكة المتحدة تبدأ (X) في الفترة من 1948 إلى 1984 . Tereno Mills حصل على نتائج الانحدار التالي(*):

$$\widehat{\Delta X_t} = 31.03 - 0.188X_{t-1}$$
$$se = (12.50) \quad (0.080)$$
$$(t =)\tau \quad (-2.35)$$

لاحظ أن: القيمة الحرجة 5% لـ τ هي 2.95− والـ 10% هي 2.60−

(a) بناء على النتائج المعطاة، هل السلسلة الزمنية الخاصة بالسكن الخاص تعتبر ساكنة أم غير ساكنة؟ بصياغة أخرى، هل يوجد جذر الوحدة لهذه السلسلة الزمنية؟ كيف عرفت ذلك؟

(b) إذا كنت ستستخدم إحصاء t التقليدي، هل القيمة المحسوبة لـ t معنوية إحصائيًا؟ على هذا الأساس هل تستنتج أن هذه السلسلة الزمنية ساكنة؟

(c) الآن دعنا نعتبر نتائج الانحدار التالية:

$$\widehat{\Delta^2 X_t} = 4.76 - 1.39\Delta X_{t-1} + 0.313\Delta^2 X_{t-1}$$
$$se = (5.06) \quad (0.236) \quad (0.163)$$
$$(t =)\tau \quad (-5.89)$$

حيث إن Δ^2 هو معامل الفرق الثاني، أي الحصول على الفرق الأول للفرق الأول. القيمة المقدرة لـ τ الآن تعتبر معنوية إحصائيًا. ما الذي يمكنك قوله الآن عن سكون السلسلة الزمنية محل الدراسة؟

لاحظ أن: الهدف من الانحدار السابق هو معرفة ما إذا كانت السلسلة الزمنية لها جذر وحدة أم لا.

24.21 قم بتوليد سلسلتي سير عشوائي كما هو موضح في (1.7.21) و (2.7.21) وقم بعمل انحدار لواحدة منهما على الثانية. كرر نفس التمرين ولكن

(*) تم تغيير بعض الرموز Trernce C. Mills, op. cit., p. 127

استخدم الآن فروقهما الأولى، واثبت أنه في هذا الانحدار قيمة R^2 تقريبًا تساوي الصفر، وإحصاء Durbin-Watson d قريب من 2.

25.21 لإثبات أن أي متغيرين، كل منهما له اتجاه عام محدد، قد يؤديان إلى انحدار زائف، حصل Charemza etal على الانحدار التالي بناء على بيانات من 30 مفردة(*):

$$\hat{Y}_t = 5.92 + 0.030X_t$$
$$t = (9.9) \quad (21.2)$$
$$R^2 = 0.92 \qquad d = 0.06$$

حيث إن $Y_1=1$، $Y_2=2$، . . . ، $Y_n=n$ و $X_1=1$، $X_2=4$، . . . ، $X_n=n^2$

(a) ما هو الاتجاه العام الموجود في Y؟ وفي X؟

(b) ارسم المتغيرين، وارسم خط الانحدار. ماهو الاستنتاج العام الذي تصل إليه من هذا الشكل البياني؟

26.21 من بيانات الفترة I–1971 إلى IV–1988 لكندا، تم الحصول على نتائج الانحدار التالي:

1. $$\widehat{\ln M1}_t = -10.2571 + 1.5975 \ln GDP_t$$
$$t = (-12.9422) \quad (25.8865)$$
$$R^2 = 0.9463 \qquad d = 0.3254$$

2. $$\widehat{\Delta \ln M1}_t = 0.0095 + 0.5833 \Delta \ln GDP_t$$
$$t = (2.4957) \quad (1.8958)$$
$$R^2 = 0.0885 \qquad d = 1.7399$$

3. $$\Delta\hat{u}_t = -0.1958\hat{u}_{t-1}$$
$$(t = \tau) \quad (-2.2521)$$
$$R^2 = 0.1118 \qquad d = 1.4767$$

حيث إن M1 = المعروض من المال M1، GDP = الناتج الكلي المحلي وكل من المتغيرين مقاس ببلايين الدولارات الكندية، ln هو اللوغاريتم الطبيعي و $\hat{u}_t$

(*) Charemza e tal., op.cit., p. 93.

تمثل البواقي المقدرة من انحدار 1 .

(a) فسر الانحدار 1 و 2 .

(b) هل تشك في أن انحدار 1 هو انحدار زائف؟ لماذا؟

(c) هل الانحدار 2 يعتبر انحداراً زائفًا؟ كيف يمكنك معرفة ذلك؟

(d) من نتائج انحدار 3، هل يتغير استنتاجك لـ b ؟ ولماذا؟

(e) الآن دعنا نعتبر الانحدار التالي:

$$\widehat{\Delta \ln \mathrm{M1}_t} = 0.0084 + 0.7340\Delta \ln \mathrm{GDP}_t - 0.0811\hat{u}_{t-1}$$
$$t = (2.0496) \quad (2.0636) \quad (-0.8537)$$
$$R^2 = 0.1066 \quad d = 1.6697$$

ماذا تستنتج من هذا الانحدار؟ هل يساعدك ذلك في معرفة ما إذا كان انحدار 1 زائف أم لا؟

26.21 الانحدار التالي معتمد على بيانات CPI للولايات المتحدة للفترة من 1960–1999، لإجمالي 40 مفردة سنوية:

1.
$$\widehat{\Delta \mathrm{CPI}_t} = 0.0372\mathrm{CPI}_{t-1}$$
$$t = (9.6427)$$
$$R^2 = 0.0304 \quad d = 0.5259 \quad \mathrm{RSS} = 203.6222$$

2.
$$\widehat{\Delta \mathrm{CPI}_t} = 1.8052 + 0.0208\mathrm{CPI}_{t-1}$$
$$t = (2.5000) \quad (2.7583)$$
$$R^2 = 0.1705 \quad d = 0.6030 \quad \mathrm{RSS} = 174.1966$$

3.
$$\widehat{\Delta \mathrm{CPI}_t} = 1.8790 + 0.5706t - 0.1158\mathrm{CPI}_{t-1}$$
$$t = (3.1460) \quad (4.2576) \quad (-3.5443)$$
$$R^2 = 0.4483 \quad d = 0.7969 \quad \mathrm{RSS} = 115.8579$$

حيث إن RSS = مجموع مربعات البواقي

(a) وفقًا للانحدارات السابقة، ماذا يمكن أن نقول عن سكون السلسلة الزمنية CPI؟

(b) كيف يمكنك الاختيار بين النماذج الثلاثة ؟

(c) المعادلة (1) هي نفسها المعادلة (3) مطروح منها الجزء الثابت المقطوع من المحور الصادي والاتجاه العام. ما هو الاختبار الذي يمكن أن تستخدمه لمعرفة ما إذا كانت القيود الضمنية للنموذج 1 متحققة أم لا؟ (**معلومة مساعدة**) استخدم اختبارات F و t لـ Dickey-Fuller .

استخدم القيم التقريبية المعطاة في ملحق D، جدول (D.7) .

الفصل الثاني والعشرون

السلاسل الزمنية في الاقتصاد القياسي . التنبؤ

TIME SERIES ECONOMETRICS: FORECASTING

سبق وأن ذكرنا في المقدمة، أن التنبؤ يعتبر جزءًا مهمًا في تحليل الاقتصاد القياسي. ولبعض الناس، يعتبر التنبؤ هو الجزء الأكثر أهمية. كيف يمكنك التنبؤ بالمتغيرات الاقتصادية، مثل GDP، التضخم، معدل تغيير العملة، أسعار الأسهم، معدلات البطالة والعديد من المتغيرات الاقتصادية التي لاحصر لها؟ في هذا الفصل، سنناقش طريقتين للتنبؤ واللتين اكتسبتا شهرة كبيرة: (1) الانحدار الذاتي والمتوسطات المتحركة التكاملية (ARIMA) والمعرفة باسم طريقة Box-Jenkins[1] . و (2) متجه الانحدار الذاتي (VAR).

في هذا الفصل، سنناقش المشاكل الخاصة بتنبؤ أسعار أصول مالية، مثل أسعار الأسهم ومعدلات تغيير العملة. أسعار الأصول مميزة بظاهرة تسمى العنقودية المتطايرة، أي فترات يكون فيها تغيير كبير لفترة زمنية ما يتبعها فترة تتميز بالهدوء النسبي. يمكنك قراءة مؤشر Dow Jones في الفترة الماضية القريبة. نماذج الانحدار الذاتي المشروط باختلاف التباين (ARCH) أو النماذج العامة للانحدار الذاتي المشروط باختلاف التباين (GARCH) توصف فيها ظاهرة مثل العنقودية المتطايرة.

موضوع التنبؤ في علم الاقتصاد، موضوع شديد الاتساع، وهناك العديد من الكتب التي تناولت هذا الموضوع بالتفصيل. هدفنا في هذا الفصل، هو إعطاء القارئ لمحة عامة عن هذا الموضوع. القارئ الذي يرغب في معرفة المزيد عن ذلك

---

(1) G.P.E. Box and G. M. Jenkins, Time Series: Forecasting and Control, revised ed., Holden Day, San Fracisco, 1978.

الموضوع عليه بالرجوع إلى قائمة المراجع لأي دراسة مستقبلية. لحسن الحظ، معظم حزم البرامج الإلكترونية الحديثة في الاقتصاد القياسي لديها مقدمة بسيطة للعديد من الأساليب التي ستتم مناقشتها في هذا الفصل.

الصلة بين هذا الفصل والفصل السابق، هي أن نماذج التنبؤ التي ستتم مناقشتها لاحقًا، تفترض أن السلسلة الزمنية محل الدراسة سلسلة ساكنة، أو من الممكن تحويلها إلى سلسلة ساكنة باستخدام التحويلة المناسبة. وكما سنرى في هذا الفصل، سيتم استخدام العديد من المصطلحات السابق التعرض لها في الفصول السابقة.

1.22 أساليب التنبؤ الاقتصادي :
APPROACHES TO ECONOMIC FORECASTING

بوجه عام، توجد خمس طرق للتنبؤ الاقتصادي المعتمد على بيانات سلسلة زمنية: (1) طرق التمهيد الأسي، (2) نماذج انحدار المعادلة المفردة، (3) نماذج انحدار المعادلات الآنية، (4) نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية (ARIMA) و (5) متجه الانحدار الذاتي.

طرق التمهيد الأسي[2]: Exponential Smoothing Methods

هذه الطرق أساسية لتوفيق المنحنى المناسب لبيانات تاريخية لسلسلة زمنية ما. هناك العديد من هذه الطرق، مثل التمهيد الأسي المفرد، طريقة Holt الخطية وطريقة Holt-Winter وتنوعاتهما. وعلى الرغم من أنهما مازالتا تستخدمان في العديد من مجالات الأعمال والتنبؤ الاقتصادي، إلا أن الطرق الأربع الأخرى المذكورة سابقًا تستخدم كثيرًا كبدائل عن هذه الطرق. لن نستعرض طرق التمهيد الأسي في هذا الفصل، حيث إن ذلك يقع خارج نطاق هذا الكتاب.

نماذج انحدار ذات المعادلة المنفردة : Single-equation regression models

معظم هذا الكتاب، يتم فيه مناقشة نماذج الانحدار ذات المعادلة الواحدة. وكمثال لنماذج المعادلة المنفردة، اعتبر دالة الطلب على السيارات من أسس النظرية الاقتصادية، نحن نفترض أن الطلب على السيارات هو دالة في أسعار السيارات،

(2) لشرح بسيط لهذه الطرق، انظر: Spyros Makridakis, Steven C. Wheelwright, and Rob J. Hyndman, Forecasting Methods and Applications, 3d ed., John Wiley & Sons, New York, 1998.

نفقات الدعاية والإعلان، دخل المستهلك، ومعدلات الفائدة (كمقياس لتكلفة الدين)، وبعض المتغيرات الأخرى ذات علاقة بالموضوع (مثل حجم الأسرة، المسافة من المنزل للعمل). من بيانات السلسلة الزمنية، نحن نقدر النموذج المناسب للطلب على السيارات (سواء نموذج خطي، خطي لوغاريتمي أو غير خطي) والذي يستخدم في التنبؤ بالطلب على السيارات في المستقبل. بالطبع كما لاحظنا في الفصل (5)، أخطاء التنبؤ تتزايد كلما بعدنا في الفترات الزمنية المتنبأ بها.

نماذج انحدار المعادلات الآنية:[3] Simultaneous- equation regression models

في الفصول 18، 19 و 20 درسنا نماذج المعادلات الآنية. من ذروتها في الفترة 1960s و1970s. وكانت النماذج الاقتصادية التفصيلية للولايات المتحدة التي اعتمدت على المعادلات الآنية استخدمت كثيراً في التنبؤ الاقتصادي. ولكن بعد ذلك انطفأ وميض تنبؤ هذه النماذج، بسبب أدائها التنبؤي الصعب، خصوصاً في مشكلة أسعار البترول في السنوات 1973 و 1979 (بسبب خطر البترول من قبل OPEC)، وأيضاً بسبب ماسمي انتقادات Lucas[4] الثقة في هذه الانتقادات، كما يمكن أن نتذكر، جاءت من القول بأن المعالم المقدرة من نموذج اقتصاد قياسي ما تعتمد على السياسة السائدة وقت تقدير النموذج، وستتغير بمجرد تغيير هذه السياسة. باختصار المعالم المقدرة غير ثابتة مع وجود تغيير في السياسة الموجودة.

فعلى سبيل المثال، في أكتوبر 1979 الهيئة الفيدرالية Fed غيرت سياستها المالية بشكل كبير، فبدلاً من استهداف معدلات الفائدة، فقد أعلنت أنها ستقوم بإجبار مرافقة معدل النمو في المعروض من المال. وعندما حدث مثل ذلك التغيير، فإن النموذج الاقتصادي المقدر من بيانات سابقة سيكون لها قدرة تنبؤية ضعيفة وفقاً للنظام الجديد. هذه الأيام الهيئة الفيدرالية أعلنت أنها بدلاً من التحكم في المعروض من المال، سيتم التحكم في معدلات الفائدة قصيرة المدى (معدل التمويل الفيدرالي).

---

(3) كتب في استخدام نماذج المعادلات الآنية للتنبؤ، انظر: Robert S. Pindyck and Daniel L. Rubin-feld, Ecomometric Models & Economtic Forecasts, 4th ed., McGraw-Hill, New York, 1998, Part III

(4) Robert E. Lucas, "Econometric policy Evaluation: A Critique, "in Carnegie-Rochester Conference Series, The Phillips Curve, North-Holland, Amsterdam, 1976, pp. 19-46
هذه المقالة من بين مقالات أخرى جعلت Lucas يحصل على جائزة نوبل في الاقتصاد.

نماذج ARIMA : ARIMA Models

المنشور في تحليل السلاسل الزمنية لـ Box و Jenkins : التنبؤ والتحكم (op. cit.) خلق جيل جديد من طرق التنبؤ. فالطريقة المعروفة باسم طريقة Box- Jenkins (BJ) ومعروفة فنيًا باسم طرق ARIMA تبرز أهمية هذه الطرق، ليس في تكوين معادلة منفردة أو نماذج معادلات آنية، ولكن في تحليل الخصائص العشوائية أو الاحتمالية لسلسلة زمنية اقتصادية وفقًا لفلسفتهم الخاصة وهي "دع البيانات تتحدث عن نفسها". فعكس نماذج الانحدار عندما تكون yt مفسرة بـ k من المتغيرات المنحدرة X_1، X_2، X_3، . . . ، X_k فإن نماذج السلاسل الزمنية من نوع BJ تسمح لـ Y_t بأن تفسر من خلال قيم الـ Y نفسها في فترات زمنية ماضية أو متأخرة، ومقدار خطأ عشوائي لهذا السبب فإن نماذج ARIMA يطلق عليها أحيانًا نماذج غير نظرية، حيث لايتم استنتاجها من أي نظرية اقتصادية. فالنظريات الاقتصادية هي عادة الأساس لنماذج المعادلات الآنية.

بشكل عابر، لاحظ أن اهتمامنا في هذا الفصل، يقتصر على نماذج ARIMA أحادية المتغير، أي نماذج ARIMA الخاصة بسلسلة زمنية منفردة. لكن التحليل يمكن أن يتوسع ويشمل نماذج ARIMA متعددة المتغيرات.

نماذج VAR : VAR Models

طريقة VAR تفترض بشكل سطحي نمذجة المعادلات الآنية، حيث نعتبر العديد من المتغيرات الداخلية معًا. ولكن كل متغير داخلي يتم تفسيره من خلال قيمة في الماضي أو فترات زمنية متأخرة، والقيم الماضية أيضًا لباقي المتغيرات الداخلية في النموذج، عادة لاتوجد متغيرات خارجية في النموذج.

في المتبقي من هذا الفصل، سنستعرض أساسيات طريقة Box -Jenkins وطريقة VAR للتنبؤ الاقتصادي. مناقشتنا ستكون ابتدائية ومساعدة على فهم الموضوع. القارئ الذي يرغب في معرفة مزيد من التفاصيل عليه اللجوء إلى قائمة المراجع[5].

(5) انظر (كتاب تطبيقي) Pindyck and Rubinfeld, op. cit., Part 3; Alan Pankratz, Forecasting with Dynamic Regression Models, John Wiley & Sons, New York, 1991

(كتاب أكثر صعوبة) Andrew Harvey, The Econometric Analysis of Time Series, The MIT Press, 2d ed., Cambridge, Mass., 1990.

مناقشة عميقة ممكنة الفهم موجودة في Terence C. Mills, Time Series Techntiques for Economists, Capbridge University Press, New York, 1990.

2.22 AR، MA و ARIMA لنمذجة بيانات السلاسل الزمنية : AR, MA, AND ARIMA MODELING OF TIME SERIES DATA

لاستعراض بعض الأفكار، بعضها قديم وبعضها جديد، دعنا نستخدم بيانات السلسلة الزمنية الخاصة بالـ GDP في الولايات المتحدة المعطاة في جدول (1.21) الرسم البياني لهذه السلسلة معطى من قبل في الشكل (1.21) (GDP بدون فروق) والشكل (9.21) (الفروق الأولى للـ GDP)، تذكر أن GDP نفسها سلسلة زمنية غير ساكنة ولكن الفروق الأولى للسلسلة الزمنية ساكنة.

إذا كانت السلسلة الزمنية ساكنة، يمكن نمذجتها بالعديد من الطرق.

عملية انحدار ذاتي (AR) : An Autoregressive (AR) process

دع Y_t تمثل GDP عند الزمن t. إذا نمذجنا Y_t كالتالي :

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + u_t \qquad (1.2.22)$$

حيث إن δ هي متوسط Y و u_t هو مقدار خطأ عشوائي غير مرتبط له توقع يساوي الصفر وتباين ثابت σ^2 (أي أنه بحت)، وبالتالي فإننا نقول إن Y_t تتبع انحداراً ذاتيًا من الدرجة الأولى أو AR(1) عملية عشوائية. وقد تناولنا ذلك بالفعل في الفصل (12). هنا قيمة Y عند الزمن t تعتمد على قيمتها في الفترة الزمنية السابقة ومقدار عشوائي، قيم Y معبر عنها بانحرافها عن وسطها الحسابي. بمعنى آخر، فإن النموذج يقول إن القيمة المتنبأ بها لـ Y عند الزمن t هي ببساطة نسبة ($= \alpha_1$) من قيمتها عند الزمن $(t - 1)$ بالإضافة إلى مقدار عشوائي عند الزمن t، ومرة أخرى، فإن قيم الـ Y معبر عنها حول قيمها المتوسطة. ولكن إذا اعتبرنا النموذج التالي :

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_3(Y_{t-2} - \delta) + u_t \qquad (2.2.22)$$

بالتالي، نحن نقول بذلك أن Y_t تتبع انحداراً ذاتيًا من الدرجة الثانية أو عملية AR(2)، بمعنى أن قيمة Y عند الزمن t تعتمد على قيمتها في الفترة السابقة بفترتين زمنيتين، قيم الـ Y معبر عنها حول وسطها الحسابي δ. عمومًا نحن لدينا

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \cdots + \alpha_p(Y_{t-p} - \delta) + u_t \qquad (3.2.22)$$

في هذه الحالة تكون Y_t تتبع انحداراً ذاتيًا من الدرجة p أو عملية AR(p). لاحظ

أنه في كل النماذج السابقة، لايوجد سوى القيم الحالية، والسابقة لـ Y لايوجد أي متغيرات منحدرة. بمعنى أن "البيانات تتحدث عن نفسها". هذا يعتبر نوعًا من النماذج مخفضة الشكل التي تناولناها في مناقشتنا لنماذج المعادلات الآنية.

عملية متوسطات متحركة (MA) : A Moving Average (MA) process

عملية AR التي ناقشناها الآن، ليست الطريقة الوحيدة لتوليد الـ Y. افترض أننا نمذجنا Y كالتالي:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} \qquad (4.2.22)$$

حيث إن μ ثابت، و u كما سبق هو مقدار خطأ عشوائي بحت. هنا Y عن الزمن t تساوي ثابتًا بالإضافة إلى متوسط متحرك للمقادير الأخطاء الحالية والماضية وبالتالي، في الوضع الحالي نقول إن Y تتبع **متوسطات متحركة من الدرجة الأولى** أو عملية **MA(1)**. ولكن إذا اتبعت Y الشكل التالي:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} \qquad (5.2.22)$$

إذن هذه عملية **MA(2)** وبشكل عام فإن:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \cdots + \beta_q u_{t-q} \qquad (6.2.22)$$

هي عملية **MA(*q*)**. باختصار عملية المتوسطات المتحركة هي ببساطة توليفة خطية من مقادير الأخطاء العشوائية البحتة.

عملية انحدار ذاتي ومتوسطات متحركة (ARMA) :

An Autoregressive and moving average (ARMA) process

بالطبع من المحتمل أن تعرف Y بكل من AR و MA، ومن هنا تأتي ARMA، أي أن Y_t تتبع عملية **ARMA (1,1)** إذًا كان يمكن كتابتها كالتالي:

$$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1} \qquad (7.2.22)$$

حيث يوجد مقدار انحدار ذاتي واحد، ومقدار متوسطات متحركة واحد. في (7.2.22) θ تمثل المقدار الثابت.

في العموم، في عملية **ARMA (*p*, *q*)** سيكون لدينا p مقادير انحدار ذاتي، و q مقادير متوسطات متحركة.

عملية انحدار ذاتي ومتوسطات متحركة مدمجة (ARIMA) :
An Autoregressive integrated moving average (ARIMA) process

نماذج السلاسل الزمنية التي ناقشناها حتى الآن، تعتمد على افتراض أن السلسلة الزمنية لديها سكون ضعيف بالمعنى الذي تم استعراضه في الفصل (21) . باختصار ، فإن المتوسط والتباين للسلسلة الزمنية ضعيفة السكون يكونان ثابتين، والتغاير لايتغير مع الزمن. ولكننا نعلم أن العديد من السلاسل الزمنية الاقتصادية غير ساكنة، وبالتالي يقال عنها مدمجة، على سبيل المثال، السلسلة الزمنية الاقتصادية في جدول (1.21) هي سلسلة مدمجة.

لكن رأينا في الفصل (21)، إنه إذا كانت السلسلة الزمنية مدمجة من الدرجة الأولى، (بمعنى $I(1)$) فإن فروقها الأولى هي $I(0)$ بمعنى أنها ساكنة. بالمثل إذا كانت السلسلة الزمنية I(2) فإن فروقها الثانية هي $I(0)$ عموماً إذا كانت السلسلة الزمنية هي $I(d)$ فإنه بعد أخذ الفروق d مرة فإننا نحصل على $I(0)$ سلسلة زمنية.

وبالتالي، إذا احتجنا إلى أخذ d من الفروق للسلسلة الزمنية لجعلها ساكنة وطبقنا نموذج ARMA (p, q) لها، فنحن نقول إن السلسلة الزمنية الأصلية هي ARIMA (p, d, q) بمعنى أنها سلسلة زمنية ذات انحدار ذاتي بمتوسطات متحركة مدمجة. حيث p ترمز إلى عدد حدود الانحدار الذاتي، d عدد المرات اللازم أخذ الفروق فيها حتى تصبح السلسلة ساكنة، و q عدد حدود المتوسطات المتحركة. وبالتالي السلسلة الزمنية ARIMA (2, 1, 2) لابد أن تأخذ لها الفروق الأولى $(d = 1)$ قبل أن تصبح ساكنة، والفروق الأولى للسلسلة الزمنية الساكنة يمكن نمذجتها كعملية ARMA (2, 2)، بمعنى أن لديها اثنين من مقادير AR واثنين من مقادير MA. بالطبع إذا كانت $d = 0$ (بمعنى أن السلسلة ساكنة منذ البداية) فإن ARIMA $(p, d = 0, q)$= ARMA (p, q) لاحظ أن ARIMA $(p, 0, 0)$ تعني عملية ساكنة لـ $AR(p)$ تامة و ARIMA $(0, 0, p)$ تعني عملية ساكنة لـ $MA(q)$ تامة.

بمعلومية قيم P، d و q، فإنه يمكن تحديد النموذج المناسب للعملية. النقطة المهمة للملاحظة عند استخدام طريقة Box-Jenkins، أننا لابد أن يكون لدينا إما سلسلة زمنية ساكنة أو سلسلة زمنية ساكنة بعد أخذ فروق أولى أو أكثر لها. السبب وراء افتراض السكون يمكن تفسيره كالتالي:

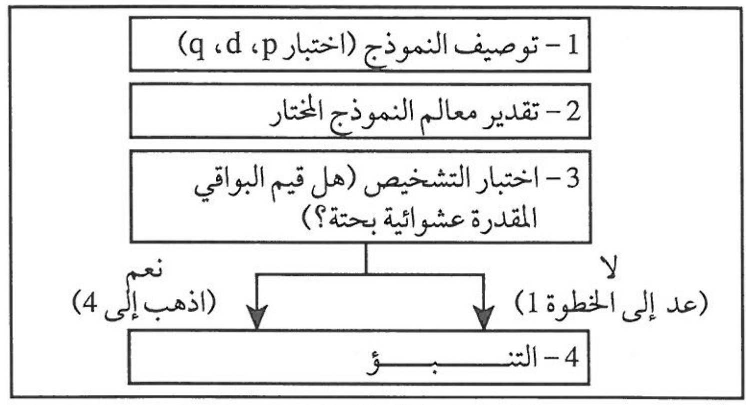
هدف B–J (Box–Jenkins) هو تعريف وتقدير النموذج الإحصائي الذي يمكنه تفسير بيانات العينة. إذا كان هذا النموذج المقدر سيستخدم بعد ذلك للتنبؤ، فلابد أن نفترض خاصية ثبات هذا النموذج مع مرور الزمن، أو عمليًا خلال فترات زمنية مستقبلية. وبالتالي السبب البسيط وراء ضرورة سكون البيانات هو أن أي نموذج يستخدم هذه البيانات يمكن أن يعرف نفسه على أنه ساكن أو مستقر، وبالتالي يعطي أساسًا سليمًا للتنبؤ[6].

3.22 طريقة Box- Jenkins (BJ):
THE BOX- JENKINS (BJ) METHODOLOGY

السؤال المهم الآن هو : بالنظر إلى السلسلة الزمنية، مثل سلسلة GDP للولايات المتحدة الموجودة في شكل (1.21). كيف يمكن معرفة ما إذا كانت تتبع عملية AR (وإذا كانت كذلك ما هي قيمة P) أو عملية MA (وإذا كانت كذلك، ما هي قيمة q) أو عملية ARMA (وإذا كانت كذلك، فما هي قيم p، q) أو عملية ARIMA وفي هذه الحالة نحتاج معرفة P، d و q. طريقة BJ تجيب عن هذا السؤال السابق. الطريقة تشتمل على أربع خطوات:

الخطوة 1- التوصيف : بمعنى تحديد القيم المناسبة لكل من P، d و q، وسنرى لاحقًا كيف يساعدنا مصور الارتباط أو مصور الارتباط الجزيئي في هذه المهمة.

الخطوة 2- التقدير : بعد تحديد القيم المناسبة لـ p، q فإن المرحلة التالية هي تقدير معالم مقادير الانحدار الذاتي والمتوسطات المتحركة الموجودة في النموذج. أحيانًا يمكن عمل هذه الحسابات بالمربعات الصغرى، وأحيانًا نحتاج إلى طريقة غير خطية (في المعلمات) في تقدير النموذج.



شكل (1.22) طريقة Box- Jenkins

(6) Michael Pokorny, An Introduction to Econometrics, Basil Blackwell, New York, 1987, p. 343.

وبما أن هذه المهمة الآن أصبح من السهل القيام بها عن طريق العديد من حزم البرامج الإلكترونية الإحصائية، لا يوجد داع للقلق من الخلفية الرياضية للتقدير، الطالب الذي يرغب في معرفة المزيد عن ذلك، عليه الرجوع إلى قائمة المراجع.

الخطوة 3- اختبار التشخيص: بعد اختيار نموذج ARIMA معين، وتقدير معالمه، نرى الآن ما إذا كان النموذج المختار يناسب البيانات بشكل جيد أم لا، حيث إنه من الممكن أن يكون هناك نموذج ARIMA آخر يقوم بذلك بشكل أفضل. ولهذا السبب، فإن نمذجة ARIMA لـ Box- Jenkins تعتبر عملية فنية أكثر منها علمية، حيث لابد من وجود بعض المهارات الخاصة لاختيار نموذج ARIMA المناسب.

أحد الاختبارات السهلة للنموذج المختار هو اختيار ما إذا كانت البواقي المقدرة من هذا النموذج عشوائية بحتة أم لا. إذا كانت كذلك فيتم قبول النموذج، وبخلاف ذلك لابد من البدء من جديد، ولذلك فطريقة BJ هي طريقة تكرارية (انظر الشكل 1.22).

الخطوة 4- التنبؤ: أحد أسباب شهرة نموذج ARIMA هو نجاحه في التنبؤ. ففي العديد من الحالات، يكون التنبؤ الذي يتم الحصول عليه من هذا النموذج أكثر صحة من نظيره الذي يتم الحصول عليه من نماذج اقتصادية أخرى تقليدية، خصوصًا المتعلقة بالتنبؤ في المدى القصير. مع مراعاة ضرورة اختبار كل حالة على حدة.

بعد هذه المناقشة العامة، دعنا نستعرض هذه الخطوات الأربع بمزيد من التفاصيل، في أثناء ذلك سنستخدم بيانات GDP المعطاة في جدول (1.21) لشرح النقاط المختلفة المتعلقة بالموضوع.

4.22 التوصيف: IDENTIFICATION

الأداة الرئيسية في التوصيف هي دالة الارتباط الذاتي (ACF)، دالة الارتباط الذاتي الجزيئي (PACF)، ومصور الارتباط الذي هو ببساطة رسم بياني لـ ACFS وPACFS مع الفترات الزمنية المتأخرة المختلفة.

في الفصل السابق عرفنا ACF (ρ_k) الخاص بالمجتمع و ACF $(\hat{\rho}_k)$ الخاص بالعينة. مفهوم الارتباط الجزيئي الذاتي يتم تعريفه من خلال مفهوم معامل الانحدار الجزيئي. في نموذج الانحدار متعدد المتغيرات الذي يوجد فيه k متغير، فإن معامل الانحدار k (β_k) يقيس معدل التغير في العينة المتوسطة للمتغير المنحدر عليه بالنسبة لتغير بمقدار الوحدة في الـ k متغير منحدر X_k بافتراض ثبات أثر باقي المتغيرات المنحدرة الأخرى.

بنفس الطريقة، فإن الارتباط الجزيئي ρ_{kk} يقيس الارتباط بين مشاهدات (السلسلة الزمنية) التي تبتعد بمقدار k فترة زمنية متأخرة بعد التحكم في الارتباط مع الفترات الزمنية المتوسطة (أي التي أقل من k). بمعنى آخر، فإن الارتباط الذاتي الجزيئي هو الارتباط بين Y_t و Y_{t-k} بعد حزف أثر الـ Y's الموجودة بينهما[7].

في فقرة 7.11، ناقشنا بالفعل مفهوم الارتباط الجزيئي في موضوع الانحدار، وأثبتنا علاقته مع الارتباطات البسيطة. مثل هذه الارتباطات الجزيئية يعتبر الآن من السهل حسابها من خلال معظم حزم البرامج الإحصائية. في شكل (2.22) نوضح مصور الارتباط، ومصور الارتباط الجزيئي للسلسلة GDP. من هذا الشكل، نجد أن هناك حقيقتين واضحتين:

أولاً: الـ ACF ينخفض ببطء شديد، كما في الشكل (8.21)، و ACF حتى الفترة الزمنية المتأخرة الـ 23 كان له معنوية إحصائيًا (لايساوي الصفر) بشكل منفرد، حيث إن جميع القيم خارج حدود فترات الثقة الـ 95%.

ثانيًا: بعد الفترة الزمنية المتأخرة الأولى، فإن PACF انخفض بشكل ملحوظ، وكل القيم لـ PACF بعد الفترة الزمنية المتأخرة الأولى غير معنوية إحصائيًا.

---

(7) في بيانات السلسلة الزمنية النسبة الكبيرة من الارتباط بين Y_t و y_{t-k} قد ترجع إلى الارتباط مع الفترات الزمنية المتأخرة الموجودة بينها مثل Y_{t-1}، Y_{t-2}، . . . ، Y_{t-k+1}. الارتباط الجزيئي يحذف تأثير هذه المتغيرات الموجودة بين Y_t و Y_{t-k}.

Lag	Sample ACF ($\hat{\rho}_k$)	Sample PACF ($\hat{\rho}_{kk}$)
1	0.969	0.969
2	0.935	−0.058
3	0.901	−0.020
4	0.866	−0.045
5	0.830	−0.024
6	0.791	−0.062
7	0.752	−0.029
8	0.713	−0.024
9	0.675	−0.009
10	0.638	−0.010
11	0.601	−0.020
12	0.565	−0.012
13	0.532	−0.020
14	0.500	−0.012
15	0.468	−0.021
16	0.437	−0.001
17	0.405	−0.041
18	0.375	−0.005
19	0.344	−0.038
20	0.313	−0.017
21	0.279	−0.066
22	0.246	−0.019
23	0.214	−0.008
24	0.182	−0.018
25	0.153	0.017

95% confidence interval — 95% confidence interval

شكل (2.22) مصور الارتباط ومصور الارتباط الجزيئي، GDP الولايات المتحدة، I–1970 إلى IV–1991

بما أن السلسلة الزمنية الخاصة بـ GDP في الولايات المتحدة، تعتبر سلسلة غير ساكنة لابد من جعلها ساكنة قبل تطبيق طريقة Box- Jenkins . في شكل (9.21) رسمنا الفروق الأولى لـ GDP. على عكس شكل (1.21) لانجد أي اتجاه عام في هذه السلسلة، مما يعني أن الفروق الأولى للسلسلة الزمنية الخاصة بـ GDP تعتبر ساكنة[8]. التطبيق المباشر لاختيار جذر الوحدة لـ Dickey- Fuller أوضح أن هذه بالضبط هي الحالة الحالية. ويمكن أيضًا رؤية ذلك من مصور الارتباط للقيم المقدرة لـ ACF وPACF المعطاة في الشكل (3.22). الآن لدينا شكل مختلف تمامًا لـ ACF و PACF. قيم الـ ACFS عند الفترات الزمنية المتأخرة 1، 8 و12 تبدو مختلفة عن الصفر إحصائيًا، تذكر من الفصل (21) أن الـ 95% فترة ثقة لـ ρ_k هي 2.089− و 2.089+ (لاحظ أنه: كما ناقشنا في الفصل (21)، حدود فترات الثقة هذه هي تقاربية،

(8) من الصعب تحديد ما إذا كان تباين هذه السلسلة ساكنة أم لا، خصوصًا حول 1979−1980 حظر البترول في 1979 والتغيرات الجوهرية في سياسات المال للهيئة الفيدرالية في 1979 قد يكون لها دور في الصعوبة التي نواجهها.

وبالتالي يمكن اعتبارها قيمًا مقربة وليست تامة). ولكن عند كل الفترات الزمنية المتأخرة الأخرى، فإنها غير مختلفة عن الصفر إحصائيًا. وهذا سليم أيضًا بالنسبة للارتباط الذاتي الجزيئي، $\hat{\rho}_{kk}$.

Lag		Sample ACF ($\hat{\rho}_k$)	Sample PACF ($\hat{\rho}_{kk}$)	
1	****	0.316	0.316	****
2	**	0.186	0.095	*
3	*	0.049	−0.038	
4	*	0.051	0.033	
5		−0.007	−0.032	
6		−0.019	−0.020	
7	*	−0.073	−0.062	*
8	****	−0.289	−0.280	****
9	*	−0.067	0.128	**
10		0.019	0.100	*
11		0.037	−0.008	
12	***	−0.239	−0.311	****
13	**	−0.117	0.011	
14	***	−0.204	−0.114	*
15	**	−0.128	−0.051	*
16		−0.035	−0.021	
17	*	−0.056	−0.019	
18		0.009	0.122	**
19	*	−0.045	−0.071	*
20	*	0.066	−0.126	**
21	*	0.084	0.089	*
22	*	0.039	−0.060	*
23	*	−0.068	−0.121	**
24		−0.032	−0.041	*
25		0.013	0.092	*

95% confidence interval 95% confidence interval

شكل (3.22) مصور الارتباط ومصور الارتباط الجزيئي للفروق الأولى لـ GDP ، الولايات المتحدة I–1970 إلى IV–1991

الآن، كيف أن مصورات الارتباط المعطاة في الشكل (3.22) تساعدنا في معرفة شكل ARMA للسلسلة الزمنية الخاصة بـ GDP؟ (لاحظ أننا: سنعتبر فقط الفروق الأولى لـ GDP حيث إنها ساكنة). أحد طرق عمل ذلك هو اعتبار ACF و PACF ومصورات الارتباط الخاصة بهما للعدد من عمليات ARMA مثل AR(1)، AR(2)، MA(1)، MA(2)، ARMA(1,1)، ARMA(2,2) وهكذا. وبما أن هذه العمليات العشوائية لها شكل محدد لـ ACF و PACF، فإذا تماثلت السلسلة الزمنية محل الدراسة مع أحد هذه الأشكال، فإننا سنعرف (نوصف) السلسلة الزمنية وفقًا لهذه

العملية. بالطبع نحتاج إلى القيام باختبار لهذا التشخيص لمعرفة ما إذا كان نموذج الـ ARMA المختار مناسبًا بشكل دقيق للبيانات، أم لا.

دراسة خصائص عمليات ARIMA المختلفة ستستعرض وقتًا ومساحة كبيرة، ولذلك فإن ما نرغب في القيام به هو إعطاء الخطوط العريضة فقط (انظر جدول 1.22)، قائمة المراجع بها المزيد من التفاصيل عن العديد من العمليات العشوائية.

لاحظ أن ACFs و PACFs لعمليات AR(q) و MA(q) لها شكل بياني عكسي، ففي حالة AR(q) فإن AC ينخفض هندسيًا أو أسيًا، ولكن PACF ينتهي بعد عدد معين من الفترات الزمنية المتأخرة، ويحدث عكس ذلك في عملية MA(q).

هندسيًا، هذه الأشكال المختلفة معطاة في شكل (4.22).

تحذير مهم. بما أننا في الواقع لا نلاحظ ACFs و PACFs النظرية، بل تعتمد على نظيرهما من العينة، فإن ACFs و PACFs المقدر لن يتماثلا بالضبط مع نظيرهما النظري. مانهتم به هو مدى التماثل بين قيم ACFs و PACFs النظرية وقيمهما من العينة، بحيث إن هذه القيم ترشدنا إلى الطريق السليم في تكوين نماذج ARIMA. ولهذا السبب فإن نمذجة ARIMA تتطلب العديد من المهارات، والتي تكتسب بطبيعة الحال بالمزيد من التجربة.

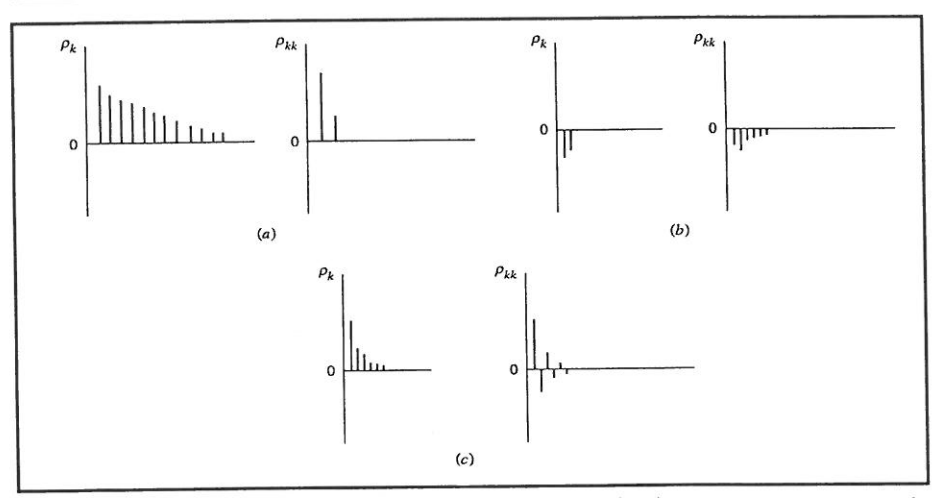
توصيف ARIMA لـ GDP الولايات المتحدة: ARIMA Identification of U.S. GDP

بالعودة إلى مصور الارتباط، ومصور الارتباط الجزيئي لـ GDP الساكن الخاص بالولايات المتحدة (بعد أخذ الفروق الأولى) للفترة 1970-I إلى 1991-IV المعطى في الشكل (3.22). ماذا ترى من ذلك؟

جدول (1.22) الشكل النظري لـ ACF و PACF

نوع النموذج	الشكل الخاص بـ ACF	الشكل الخاص بـ PACF
AR(p)	ينخفض أسيًا أو بشكل تموجي أو الاثنين معًا	قفزات معنوية خلال الفترة الزمنية المتأخرة p
MA(q)	قفزات معنوية خلال الفترة الزمنية المتأخرة q	ينخفض أسيًا
ARMA(p, q)	انخفاض جزيئي	انخفاض أسي

لاحظ أن: المصطلح انخفاض هندسي أو أسي، يعني نفس الشئ (تذكر مناقشتنا لنموذج Koyck المتوزع متأخرًا).



شكل (4.22) ACF و PACF لعمليات عشوائية معنوية: (a) AR(2): $\alpha_1 = 0.5$، $\alpha_2 = 0.3$
(b) MA(2): $\beta_1 = 0.5$، $\beta_2 = 0.3$ (c) ARMA(1.1): $\alpha_1 = 0.5$، $\beta_1 = 0.5$

تذكر أن ACF و PACF الموضحة هنا، هي كميات من العينة، ولايوجد لدينا شكل واضح كالمقترح في جدول (1.22). الارتباط الذاتي ينخفض حتى الفترة المتأخرة 4. وبعد ذلك بخلاف الفترات الزمنية المتأخرة 8، 12 فإن المتبقي منها غير مختلفة عن الصفر إحصائيًا (الخطوط المتصلة الموجودة في هذا الشكل تعطي حدود فترات الثقة الـ 95% التقريبية). الارتباط الذاتي الجزيئي مع الفقرات الموجودة في الفترات الزمنية المتأخرة 1، 8 و12 يبدو معنويًا إحصائيًا، ولكن الباقي غير معنوي، إذا كان معامل الارتباط الجزيئي معنويًا فقط عند الفترة الزمنية المتأخرة 1، فإننا نستنتج إمكانية التوصيف كنموذج AR(1) وبالتالي دعنا نفترض أن العملية المولدة للفروق الأولى لـ GDP على الأكثر ستكون عملية AR(12). بالطبع ليس بالضرورة أن يحتوي النموذج على كل مقادير AR حتى الـ 12، حيث إننا من مصور الارتباط الجزيئي نعرف أن مقادير AR المعنوية هي 1، 8 و 12 فقط.

5.22 تقدير نموذج ARIMA :

ESTIMATION OF THE ARIMA MODEL

دع Y_t^* ترمز إلى الفروق الأولى لـ GDP في الولايات المتحدة. وبالتالي، فإن نموذج AR هو:

$$Y_t^* = \delta + \alpha_1 Y_{t-1}^* + \alpha_8 Y_{t-8}^* + \alpha_{12} Y_{t-12}^* \qquad (1.5.22)$$

باستخدم Eviews نحصل على المقدرات التالية:

$$\begin{aligned}
\widehat{Y_t^*} &= 23.0894 + 0.3428Y_{t-1}^* - 0.2994Y_{t-8}^* - 0.2644Y_{t-12}^* \\
\text{se} &= (2.9774) \quad (0.0987) \quad (0.1016) \quad (0.0986) \\
t &= (7.7547) \quad (3.4695) \quad (-2.9475) \quad (-2.6817) \\
& R^2 = 0.2931 \qquad d = 1.7663
\end{aligned} \tag{2.5.22}$$

سنترك للقارئ تقدير النموذج الذي يحتوي فقط على Y_{t-1}^* كتمرين، وأيضًا تقدير نموذج آخر يحتوي على كل من Y_{t-1}^* و Y_{t-8}^* ومقارنة هذين النموذجين مع النتائج المعطاة في (2.5.22).

6.22 اختبار التشخيص : DIAGNOSTIC CHECKING

كيف يمكن أن نعرف أن النموذج في (2.5.22) مناسب لتوفيق البيانات؟ أبسط طرق التشخيص هي الحصول على البواقي من (2.5.22) ونحصل على الـ ACF والـ PACF لهذه البواقي حتى الفترة الزمنية المتأخرة الـ 25. قيم AC، PACF المقدرة معطاة في الشكل (5.22). وكما يوضح هذا الشكل، لايوجد أي ارتباط ذاتي، أو ارتباط ذاتي جزيئي معنوي إحصائيًا بشكل منفرد.

Autocorrelations	Partial autocorrelations	Lag	ACF ($\hat{\rho}_k$)	PACF ($\hat{\rho}_{kk}$)
*	*	1	0.043	0.043
*	*	2	0.113	0.112
		3	0.020	0.012
*	**	4	−0.100	−0.116
*	*	5	−0.068	−0.065
		6	−0.029	0.001
*		7	−0.040	−0.019
*	**	8	−0.112	−0.118
*	*	9	0.065	0.069
**	**	10	0.126	0.151
*	*	11	0.099	0.076
	*	12	−0.026	−0.106
**	*	13	0.120	0.102
**	**	14	−0.181	−0.150
**	**	15	−0.128	−0.131
*	*	16	−0.073	−0.050
**		17	−0.121	−0.038
	*	18	0.017	0.059
		19	−0.007	−0.027
*	**	20	−0.085	−0.163
*	*	21	0.055	0.059
		22	0.010	−0.016
	*	23	−0.038	−0.103
*	*	24	−0.053	−0.072
	*	25	−0.002	0.100

95% confidence interval — 95% confidence interval

Box–Pierce Q statistic 14.42 Probability 0.9540 se of correlations 0.110

Ljung–Box (LB) Statistic 17.63 Probability 0.8578

شكل (5.22) مصور الارتباط للبواقي التي تم الحصول عليها من نموذج (2.5.22) ARIMA

حتى مجموع مربعات الـ 25 ارتباطًا ذاتيًا، كما موضح من إحصاءات Box-Pierce Q و Ljung- Box LB (انظر الفصل 21)، غير معنوية إحصائيًا. بمعنى آخر، فإن مصور الارتباط سواء للارتباط الذاتي أو الارتباط الذاتي الجزيئي يعطي الانطباع بأن البواقي المقدرة من (2.5.22) تامة العشوائية. وبالتالي لاتوجد ضرورة للبحث عن نموذج ARIMA آخر.

7.22 التنبؤ : FORECASTING

بالعودة إلى بيانات GDP للفترة من I–1970 إلى IV–1999. باستخدام النموذج (2.5.22) نريد أن نتنبأ بالـ GDP للفترات الربع سنوية الأربعة الأولى لسنة 1992. ولكن في (2.5.22) المتغير التابع هو التغير في الـ GDP خلال الفترة الربع سنوية السابقة. وبالتالي، إذا استخدمنا (2.5.22) فالذي نحصل عليه هو التنبؤ بتغيرات الـ GDP بين الربع الأول لـ 1992 والربع الرابع لـ 1991، الربع الثاني لـ 1992 والربع الأول لـ 1992 وهكذا.

للتنبؤ بمستوى الـ GDP بدلاً من التغير فيه، يمكن ألا نستخدم تحويلة الفروق الأولى التي استخدمناها للحصول على التغيرات. (بمعنى أكثر فنية، نقوم بعمل تكامل لسلسلة الفروق الأولى)، وبالتالي للحصول على قيمة متنبأة لـ GDP (بدلاً من ΔGDP) للفترة I–1992، نعيد كتابة النموذج (1.5.22) كالتالي:

$$
\begin{aligned}
Y_{1992-\text{I}} - Y_{1991-\text{IV}} = \delta + \alpha_1[Y_{1991-\text{IV}} - Y_{1991-\text{III}}] \\
+ \alpha_8[Y_{1989-\text{IV}} - Y_{1989-\text{III}}] \\
+ \alpha_{12}[Y_{1988-\text{IV}} - Y_{1988-\text{III}}] + u_{1992-\text{I}} \qquad \textbf{(1.7.22)}
\end{aligned}
$$

وهذا هو

$$
\begin{aligned}
Y_{1992-\text{I}} = \delta + (1+\alpha_1)Y_{1991-\text{IV}} - \alpha_1 Y_{1991-\text{III}} + \alpha_8 Y_{1989-\text{IV}} - \alpha_8 Y_{1989-\text{III}} \\
+ \alpha_{12} Y_{1988-\text{IV}} - \alpha_{12} Y_{1988-\text{III}} + u_{1992-\text{I}} \qquad \textbf{(2.7.22)}
\end{aligned}
$$

القيم δ، α_1، α_8 و α_{12} معروفة بالفعل من الانحدار المقدر (2.5.22) قيمة M1992-I مفترض أن تساوي الصفر (لماذا؟). وبالتالي يمكن بسهولة الحصول على القيم التنبؤية لـ $Y_{1992-\text{I}}$. التقدير الرقمي لهذه القيم المتنبأ بها هو[9].

(9) على الرغم من أن حزم البرامج الإلكترونية تقوم بكل هذه الحسابات بشكل عادي، فإننا وضحنا تفاصيل هذه الحسابات لشرح الطريقة التي تتضمن كل هذه الحسابات.

$$\begin{aligned}
\hat{Y}_{1992-\text{I}} &= 23.0894 + (1 + 0.3428)Y_{1991-\text{IV}} - 0.3428Y_{1991-\text{III}} \\
&\quad + (-0.2994)Y_{1989-\text{IV}} - (-0.2994)Y_{1989-\text{III}} \\
&\quad + (-0.2644)Y_{1988-\text{IV}} - (-0.2644)Y_{1988-\text{III}} \\
&= 23.0894 + 1.3428(4868) - 0.3428(4862.7) \\
&\quad - 0.2994(4859.7) + 0.2994(4845.6) - 0.2644(4779.7) \\
&\quad + 0.2644(4734.5) \\
&= 4876.7 \qquad \text{(approx.)}
\end{aligned}$$

وبالتالي القيم المتنبأ بها لـ GDP في I–1992 هي حوالي 4877 بليون دولار (دولارات 1987). القيمة الحقيقية للـ GDP الحقيقي في I–1992 كان 4873,7 $ بليون دولار. وبالتالي الخطأ في التنبؤ كان تقديراً زائداً عن القيمة الحقيقية بـ 3 بلايين دولار.

لاحظ أنه إذا استخدمت (2.5.22) لحساب التنبؤ بالتغير في الـ GDP خلال الفترة الزمنية IV–1991 إلى I–1992 ستحصل على 4.25– بليون دولار.

8.22 جوانب أخرى لطريقة BJ : FURTHER ASPECTS OF THE BJ METHODOLOGY

في الفقرات السابقة، قدمنا مقدمة مبدئية لنمذجة BJ. هناك العديد من الجوانب الأخرى لهذه الطريقة، والتي لم نتطرق إليها نظراً لضيق المساحة. على سبيل المثال، الموسمية. العديد من السلاسل الزمنية يظهر فيها مايسمى بالسلوك الموسمي. كأمثلة على ذلك: المبيعات الخاصة بمحل ما وعلاقة ذلك بالإجازات الرئيسية، الاستهلاك الموسمي للآيس كريم، السفر أثناء العطلات العامة الرسمية وهكذا. إذا أخذنا على سبيل المثال بيانات المبيعات لمحل ما الربع سنوية، فالشكل البياني للمبيعات سيوضح قفزات في الربع الأخير من السنة (الربع الرابع). في مثل هذه المواقف، يمكن إزالة أثر الموسم بأخذ الفروق في الربع الرابع للمبيعات، ثم تحديد بعد ذلك أي نوع من نماذج ARIMA يناسب البيانات أكثر. قد قمنا بتحليل سلسلة زمنية واحدة فقط. ولكن لايوجد مايمنع من تطبيق طريقة BJ على دراسة آنية أو أكثر من السلاسل الزمنية.

التعمق في مثل هذه النقطة يقع خارج نطاق هذا الكتاب. القارئ الذي يرغب في معرفة المزيد عليه باللجوء إلى قائمة المراجع[10]. في الفقرة التالية سنناقش هذا الموضوع، ولكن في إطار ما يسمى متجه الانحدار الذاتي.

(10) لقراءة سهلة في هذا الموضوع، انظر Terence C. Mills, op. cit., part III

9.22 متجه الانحدار الذاتي (VAR) : VECTOR AUTOREGRESSION (VAR)

من الفصل (18) إلى الفصل 20، درسنا نماذج المعادلات الآنية أو البنائية. في مثل هذه النماذج، بعض المتغيرات يتم التعامل معها على إنها متغيرات داخلية، والبعض الآخر كمتغيرات خارجية أو محددة سابقًا (خارجية وداخلية في فترات زمنية متأخرة). قبل تقدير مثل هذه النماذج، لابد أن نتأكد من أن المعادلات الموجودة في النظام يمكن توصيفها (سواء تامة التوصيف أو موصفة بأكثر مما يجب). هذا التوصيف يتم الحصول عليه بافتراض أن بعض المتغيرات المحددة سابقًا تظهر فقط في بعض المعادلات. هذا القرار عادة يعتبر قرارًا شخصيًا وتم انتقاده بشكل حاد من قبل Christopher Sims[11].

وفقًا لنقد Sims، إذا كانت هناك آنية حقيقية بين مجموعة من المتغيرات، لابد من معاملتها بقدر متساو، فلابد ألا تكون هناك تفرقة مسبقة بين متغيرات داخلية وأخرى خارجية. وفقًا لهذه الفكرة قام Sims بعمل نموذج VAR.

أساس هذا النموذج تم التعرض له من قبل في اختبار السببية لـ Granger الذي ناقشناه في الفصل (17). في المعادلات (1.14.17) و (2.14.17)، التي يتم فيها تفسير الـ GDP الحالي في ضوء القيم الزمنية المتأخرة للمعروض من المال، والقيم الزمنية المتأخرة للـ GDP وتم فيها أيضًا تفسير المعروض حاليًا من المال في ضوء المعروض من المال في فترات زمنية متأخرة، والقيم المتأخرة زمنيًا للـ GDP، تم بشكل رئيسي معاملة كل من الـ GDP والمعروض من المال كزوج من المتغيرات الداخلية. ولايوجد أي متغيرات خارجية في هذا النظام.

بالمثل في المثال 13.17 درسنا العلاقة السببية بين المال ومعدل الفائدة في كندا في معادلة المال، تظهر قيم المال في فترات زمنية متأخرة قيم ومعدلات الفائدة فقط، وفي معادلة معدل الفائدة، تظهر قيم معدلات الفائدة في فترات زمنية متأخرة وقيم المال فقط.

كل من هذه الأمثلة، تعتبر توضيحًا لنماذج متجه الانحدار الذاتي، فمصطلح الانحدار الذاتي، يرجع إلى ظهور قيم للمتغير التابع في فترات زمنية متأخرة على الجانب الأيمن من المعادلة، ومصطلح متجه يرجع إلى حقيقة تعاملنا مع متجه من متغيرين أو أكثر.

---

(11) C.A. Sims, "Macroeconomics and Reality", Econometric, vol. 48, 1980, pp. I-48.

تقدير VAR : Estimation of VAR

بالعودة إلى مثال المال- معدل الفائدة الكندي، رأينا أنه عندما أدخلنا ست فترات زمنية متأخرة لكل متغير كمتغيرات منحدرة، لم نستطع رفض الفرض الخاص بوجود علاقة سببية مزدوجة بين المال (M_1) ومعدل الفائدة، R (معدل فائدة في 90 يومًا). بمعنى أن M_1 تؤثر على R و R تؤثر على M_1. مثل هذه المواقف تعتبر مثالية لتطبيق VAR عليها.

لشرح كيفية تقدير VAR، سنستخدم نفس المثال السابق. حتى الآن فرضنا أن كل معادلة تحتوي على k فترات زمنية متأخرة لـ M (مقاس كـ M_1) و k في مثل هذه الحالة، يمكن تقدير كل من المعادلات الآنية باستخدام OLS[12].

$$M_{1t} = \alpha + \sum_{j=1}^{k} \beta_j M_{t-j} + \sum_{j=1}^{k} \gamma_j R_{t-j} + u_{1t} \qquad (1.9.22)$$

$$R_t = \alpha' + \sum_{j=1}^{k} \theta_j M_{t-j} + \sum_{j=1}^{k} \gamma_j R_{t-j} + u_{2t} \qquad (2.9.22)$$

حيث إن u's هو مقادير الأخطاء العشوائية، وتسمى دوافع أو ابتكارات أو صدمات في إطار نماذج VAR.

قبل البدء في تقدير (1.9.22) و (2.9.22)، لابد أن نقرر طول الفترات الزمنية المتأخرة. هذا هو سؤال عملي. فنحن لدينا مجموع 40 مفردة، وبالتالي وجود الكثير من المقادير في فترات زمنية متأخرة سيستهلك درجات حرية عديدة، بالإضافة إلى زيادة إمكانية ظهور مشكلة الارتباط المتعدد.

وأيضًا وجود القليل من الفترات الزمنية المتأخرة سيؤدي إلى أخطاء في التوصيف (التعريف). أحد الطرق المناسبة للإجابة عن السؤال السابق، هو استخدام طريقة مثل Akaike أو Schwarz ونختار النموذج الذي يعطي أقل القيم لهذه الطرق. ولايوجد شك في أهمية استخدام فكرة التجربة والخطأ.

(12) يمكن استخدام طريقة SURE (الانحدار غير المرتبط ظاهريًا) لتقدير كل معادلتين معًا. عمومًا بما أن كل انحدار يحتوي على نفس العدد من المتغيرات الداخلية المتأخرة، فإن تقدير OLS لكل معادلة على حدة سينتج عنه تقديرات جيدة (وكافية).

لشرح هذه الطريقة، سنستخدم مبدئيًا أربع فترات زمنية متأخرة ($k = 4$) لكل متغير وباستخدام Eviews 4 حصلنا على مقدرات معالم المعادلتين السابقتين، وهذه التقديرات معطاة في جدول (2.22).

جدول (2.22) تقديرات متجه الانحدار الذاتي لأربع فترات زمنية متأخرة
العينة (معدلة) : I–1980 إلى IV–1987
المشاهدات المتضمنة في العينة : 32 بعد تعديل نقطة النهاية الأخطاء القياسية في ( ) وإحصاء t في [ ]

	M_1	R
M_1 (−1)	1.076737 (0.20174) [5.33733]	0.001282 (0.00067) [1.90083]
M_1 (−2)	0.173433 (0.31444) [0.55157]	−0.002140 (0.00105) [−2.03584]
M_1 (−3)	−0.366465 (0.34687) [−1.05648]	0.002176 (0.00116) [1.87699]
M_1 (−4)	0.077602 (0.20789) [0.37329]	−0.001479 (0.00069) [−2.12855]
R (−1)	−275.0293 (57.2174) [−4.80675]	1.139310 (0.19127) [5.95670]
R (−2)	227.1750 (95.3947) [2.38142]	−0.309053 (0.31888) [−0.96917]
R (−3)	8.511851 (96.9176) [0.08783]	0.052361 (0.32397) [0.16162]
R (−4)	−50.19926 (64.7554) [−0.77521]	0.001076 (0.21646) [0.00497]
C	2413.827 (1622.65) [1.48759]	4.919000 (5.42416) [0.90687]
R^2	0.988154	0.852890
Adj. R^2	0.984034	0.801721
Sum square residuals	4820241.	53.86233
SE equation	457.7944	1.530307
F statistic	239.8315	16.66815
Log likelihood	−236.1676	−53.73716
Akaike A/C	15.32298	3.921073
Schwarz SC	15.73521	4.333311
Mean dependent	28514.53	11.67292
SD dependent	3623.058	3.436688
Determinant residual covariance	490782.3	
Log likelihood (df adjusted)	−300.4722	
Akaike information criterion	19.90451	
Schwarz criterion	20.72899	

لاحظ أنه على الرغم من أن العينة تبدأ من I–1979 إلى 4–1988، فإننا استخدمنا العينة للفترة I–1979 إلى 4–1987، واحتفظنا بالمشاهدات الأربع الأخيرة للتأكد من دقة تنبؤنا باستخدام VAR.

بما أن المعادلات السابقة هي انحدارات OLS، فإن نتائج الانحدار المعطاة في جدول (2.22) يمكن تفسيرها بالشكل التقليدي العادي. بالطبع مع وجود العديد من قيم نفس المتغيرات في فترات زمنية متأخرة، فإن كل معامل مقدر لن يكون معنويًا

إحصائيًا، وذلك محتمل بسبب مشكلة الارتباط المتعدد. ولكن بشكل تجميعي، فإنها جميعًا قد تكون معنونة معًا على أساس قيمة اختبار F القياسي.

دعنا نستعرض النتائج المقدمة في جدول (2.22). أولاً دعنا نتناول انحدار M_1 نجد أن M_1 في الفترة الزمنية المتأخرة 1 و R عند الفترة الزمنية المتأخرة 1 و 2 لهما معنوية إحصائية منفردين. ولكن قيمة F كبيرة جداً، بحيث لا تستطيع رفض فرض أن كل القيم المتأخرة زمنيًا مجتمعة معًا لها معنوية إحصائية. أما بالنسبة لانحدار معدل الفائدة، نرى أن كل القيم المتأخرة الأربع للمال لها معنوية إحصائية منفردة (عند 10% أو أفضل) في حين أن الفترة الزمنية المتأخرة الأولى فقط لمتغير معدل الفائدة هي الوحيدة المعنوية إحصائيًا.

بغرض المقارنة، فإننا نقدم في جدول (3.22) نتائج VAR على أساس فترتين زمنيتين متأخرتين فقط لكل من المتغيرات الداخلية. سنرى هنا أنه في انحدار المال قيمة متغير المال في الفترة الزمنية المتأخرة الأولى وكل من قيم معدلات الفائدة في الفترتين الزمنيتين المتأخرتين له معنوية إحصائية منفردة، أما بالنسبة لانحدار معدل الفائدة، فإن كلاً من مقادير المال المتأخرة (عند حوالي 5%) وقيمة واحدة متأخرة لمعدل الفائدة لها معنوية إحصائية منفردة.

إذا كنا بصدد الاختيار بين النموذج المعطى في جدول (2.22) والمعطى في جدول (3.22)، من سنختار؟ قيم معلومات Akaike و Schwarz للنموذج الموجود في جدول (2.22) هي بالترتيب 15.33 و 15.73 في حين القيم المناظرة في جدول (3.22) هي 15.1 و 15.33. وبما أن أقل قيمة لإحصاء Akaike و Schwarz تعني نموذجًا أفضل، فإن النموذج الذي يصلح أكثر هو الموجود، في جدول (3.22). وعندما اعتبرنا 6 فترات زمنية متأخرة لكل المتغيرات الداخلية الموجودة وجدنا أن قيم إحصاءات Akaike و Schwarz هي 15.37 و 15.98 بالترتيب. مما يؤيد مرة أخرى اختيار النموذج ذي الفترتين الزمنيتين المتأخرتين لكل متغير داخلي، أي النموذج الموجود في جدول (3.22).

جدول (3.22) تقديرات متجه الانحدار الذاتي باستخدام فترتين زمنيتين متأخرتين
العينة (معدلة) : 1979-I إلى 1987-IV
المشاهدات المتضمنة في العينة : 34 بعد تعديل نقطة النهاية الأخطاء القياسية في ( ) وإحصاء t في [ ]

	M_1	R
M_1 (−1)	1.037537 (0.16048) [6.46509]	0.001091 (0.00059) [1.85825]
M_1 (−2)	−0.044661 (0.15591) [−0.28646]	−0.001255 (0.00057) [−2.19871]
R (−1)	−234.8850 (45.5224) [−5.15977]	1.069081 (0.16660) [6.41708]
R (−2)	160.1560 (48.5283) [3.30026]	−0.223364 (0.17760) [−1.25768]
C	1451.977 (1185.59) [1.22468]	5.796434 (4.33894) [1.33591]
R^2	0.988198	0.806660
Adj. R^2	0.986571	0.779993
Sum square residuals	5373510.	71.97054
SE equation	430.4573	1.575355
F statistic	607.0720	30.24878
Log likelihood	−251.7446	−60.99215
Akaike A/C	15.10263	3.881891
Schwarz SC	15.32709	4.106356
Mean dependent	28216.26	11.75049
SD dependent	3714.506	3.358613
Determinant residual covariance	458485.4	
Log likelihood (df adjusted)	−318.0944	
Akaike information criterion	19.29967	
Schwarz criterion	19.74860	

التنبؤ باستخدم VAR : Forecasting with VAR

افترض أننا اخترنا النموذج المعطى في جدول (3.22). يمكننا استخدامه بغرض التنبؤ بقيم M_1 و R. تذكر أن البيانات متاحة في الفترة 1979-I إلى 1988-IV ولكن لم تستخدم قيم سنة 1988 في تقدير نماذج VAR. الآن افترض أننا نريد التنبؤ بقيمة M1 لـ 1988-I، أي في الربع الأول من 1988. القيمة التنبؤية لـ 1988-I يمكن الحصول عليها كالتالي:

$$\hat{M}_{1988\text{-I}} = 1451.977 + 1.0375 M_{1987\text{-IV}} - 0.0446 M_{1987\text{-III}}$$
$$- 234.8850 R_{1987\text{-IV}} + 160.1560 R_{1987\text{-III}}$$

حيث قيم المعاملات تم الحصول عليها من جدول (3.22). والآن باستخدام القيم المناسبة لـ M و R من جدول (3.17). القيمة التنبؤية للمال في الربع الأول من 1988 تساوي 36.996 (مليون دولار كندي). القيمة الفعلية لـ M في 1988-I كانت 480 و 36، مما يعني أن نموذجنا قدر القيمة بزيادة حوالي 516 (مليون دولار) عن القيمة الحقيقية، أي حوالي 1.4% من القيمة الفعلية في 1988-I. بالطبع هذه

التقديرات ستتغير وفقًا لعدد قيم الفترات الزمنية المتأخرة التي ستوجد في نموذج VAR. متروك للقارئ أن يتنبأ بقيمة R في الربع الأول من 1988 كتمرين، ثم يقارن بين القيمة التنبؤية والقيمة الفعلية في هذا الربع الأول.

Var والسببية : VAR and Causality

ناقشنا من قبل موضوع السببية في الفصل (17). حيث درسنا اختيارات السببية لـ Sims's و Granger's. هل هناك أي علاقة بين VAR والسببية؟ في الفصل (17) (الفقرة 14) رأينا أنه حتى الفترات الزمنية المتأخرة 2، 4 و 6 كانت هناك سببية مزدوجة بين M_1 و R ولكن عند الفترة الزمنية المتأخرة 8 لم تكن هناك أي سببية مزدوجة بين المتغيرين. وبالتالي فالنتائج مخلوطة. الآن دعنا نسترجع نظرية التمثيل لـ Granger والتي ناقشناها في الفصل (21). إحدى نتائج هذه النظرية هي أنه إذا كان هناك متغيران مثلاً X_t و Y_t بينهما اندماج مزدوج وكل منهما $I(1)$ على حدة أي مدمجة من الدرجة 1 (بمعنى أن كلاً منهما على حدة غير ساكنة) فإن إما X_t لها سببية Granger لـ Y_t أو Y_t لها سببية- Granger لـ X_t.

في مثالنا التوضيحي، فإن ذلك يعني إذا كانت M_1 و R كل منهما $I(1)$ على حدة ولكن بينهما اندماج مزدوج فإنه إما M_1 لها سببية- Granger لـ R أو R لها سببية- Granger لـ M_1. مما يعني أننا يجب أولاً أن نعرف ما إذا كان هذان المتغيران $I(1)$ منفردين أم لا، ثم بعد ذلك، يجب أن نعرف ما إذا كان بينهما اندماج مزدوج أم لا. إذا لم تكن تلك هي الحالة، فإن سؤال السببية بأكمله قد يصبح سؤالاً جدليًا. في تمرين 22.22 يسأل القارئ عن معرفة ما إذا كان المتغيران غير ساكنين أم لا، ولكن باعتبار أن بينهما اندماج مزدوج. إذا قمت بحل التمرين، ستجد أن هناك دليلاً ضعيفًا على الاندماج المزدوج بين M_1 و R وذلك هو السبب وراء اللبس الحادث في اختبارات السببية التي تمت مناقشتها في الفقرة 14.17.

بعض مشاكل نمذجة VAR : Some problems with VAR modeling

المؤيد لـ VAR يرى الفضائل التالية لهذه الطريقة :

1 - طريقة بسيطة، حيث لاداعي للقلق بشأن تحديد أي المتغيرات داخلية وأيها خارجية. فكل المتغيرات في VAR داخلية[13].

(13) أحيانًا يتم تضمين متغيرات خارجية تمامًا للنموذج لدراسة الاتجاه العام والعوامل الموسمية.

2 – التقدير البسيط، حيث إنه يتم بطريقة OLS التقليدية، والتي يمكن تطبيقها لكل معادلة على حدة.

3 – التنبؤ الذي يتم الحصول عليه بهذه الطريقة في أغلب الحالات يكون أفضل من نظيره الذي يتم الحصول عليه من نماذج معادلات آنية أكثر تعقيداً[14].

ولكن الناقد لنمذجة VAR يشير إلى المشاكل التالية:

1 – على عكس نمذجة المعادلات الآنية، فإن نموذج VAR يعتبر نموذجًا نظريًا، حيث يستخدم بمعلومات سابقة أقل. تذكر أنه في نماذج المعادلات الآنية إدخال أو استبعاد متغيرات ما يلعب دوراً مهمًا في توصيف النموذج.

2 – بسبب تأكيد نموذج VAR على التنبؤ، يعتبر النموذج الأقل مناسبة لتحليل السياسات.

3 – التحدي العملي الأكبر لنمذجة VAR هو اختيار الطول المناسب للفترات الزمنية المتأخرة. افترض أن لدينا ثلاثة متغيرات في نموذج VAR وقد اخترت وجود ثمان فترات زمنية متأخرة لكل متغير في كل معادلة. بالتالي عليك تضمين 24 معلمة لفترات زمنية متأخرة في كل معادلة، بالإضافة إلى مقدار ثابت، أي مجموع 25 معلمة. إذا لم يكن حجم العينة كبيراً سيكون تقدير هذا الكم من المعلمات مستهلكًا لكثير من درجات الحرية، وذلك يتبعه العديد من المشاكل[15].

4 – وبشكل ضروري، فإنه في نموذج VAR التي يشتمل على m متغير، فإن كل الـ m متغير لابد أن يكونوا ساكنين (بشكل مشترك). إذا لم تكن هذه هي الحالة، سنحتاج إلى تحويل البيانات بشكل مناسب (مثلاً بالحصول على الفروق الأولى)

(14) انظر على سبيل المثال، T.Kindal and J. B. Ratner, "Regional Forecasting Models with Vector Autoregression: The Case of New York State, "Discussion Paper #155, Department of Economics, State University of New York at Albany, 1982.

(15) إذا كان لدينا m معادلة في نموذج VAR، و p قيمة في فترات زمنية متأخرة للـ m متغير، فإن علينا تقدير $(m + pm^2)$ معلمة ككل.

وكما لاحظ Harvey فإن النتائج المستخلصة من البيانات المحولة قد تكون غير كافية. ولاحظ أيضًا أن "الطريقة العادية المعتمدة على VAR aficio nados تعمل في مستويات حتى إذا كانت بعض هذه السلاسل غير ساكنة. في هذه الحالة، من المهم معرفة تأثير جذر الوحدة على توزيع المقدرات"[16]. والأسوأ من ذلك هو أنه إذا احتوى النموذج على مزيج من متغيرات $I(0)$ و $I(1)$ فإن هذا المزج بين المتغيرات الساكنة وغير الساكنة سيجعل تحويل البيانات أمرًا غير سهل.

5 - بما أن المعاملات المنفردة في نماذج VAR المقدرة غالبًا صعبة التفسير، فإن المستخدمين لهذا الأسلوب عادة يقدرون ما يسمى دالة استجابة الدافع (IRF). الـ IRF توضح استجابة المتغير التابع في نظام VAR للصدمات الموجودة في مقادير الأخطاء، مثل u_1 و u_2 في المعادلات (1.9.22) و (2.9.22). افترض أن u_1 في معادلة M_1 يزداد بقيمة واحدة عن الانحراف المعياري. مثل هذا التغير سيغير M_1 في الوقت الحالي، كما سيغيره في الفترات المقبلة. ولكن بما أن u_1 يظهر في انحدار R، فإن التغير في u_1 سيكون له تأثير على R. بالمثل عندما يتغير u_2 بوحدة انحراف معياري في معادلة R سيكون لذلك تأثير على u_1. الـ IRF توضح تأثير مثل هذه الصدمات أو التغيرات خلال عدة فترات في المستقبل. وعلى الرغم من أن المنفعة التي سنحصل عليها من تحليل IRF هي محل تساؤل من قبل الباحثين، إلا أنه خطوة رئيسية في تحليل VAR[17].

للمقارنة بين أداء VAR مع طرق تنبؤ أخرى، على القارئ اللجوء إلى قائمة المراجع المعطاة[18].

---

(16)Andrew Harvey, The Econometric Analysis of Time Series, The MIT Press, 2d ed., Cambridge, Mass., 1990. P. 83.

(17) D. E, Runkle, "Vector Autoregression and Reality," Journal of Business and Economic Statistics, vol. 5. 1987, pp. 437–454.

(18) S. McNees, "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts, "Journal of Business and Economic Statistics, vol. 4. 1986, pp. 5–15; E. Mahmoud, "Accuracy in Forecasting: A Survey, "Journal of Forecasting, vol. 3, 1984, pp. 139–159.

تطبيق على VAR: نموذج VAR للاقتصاد في تكساس:

An application of VAR: VAR model of the Texas economy

لاختبار المقولة الشهيرة، "كلما زادت أسعار البترول كلما تحسن الاقتصاد في تكساس". قام Thomas Fomby و Joseph Hirschberg بعمل نموذج VAR له ثلاثة متغيرات للاقتصاد في تكساس للفترة 1974–I إلى 1988–I [19] المتغيرات الثلاثة هي:

1 – نسبة التغير في السعر الحقيقي للبترول.

2 – نسبة التغير في الوظائف غير الزراعية في تكساس.

3 – نسبة التغير في الوظائف غير الزراعية في باقي الولايات المتحدة.

الباحثون استخدموا مقداراً ثابتًا وفترتين زمنيتين متأخرتين لكل متغير في كل معادلة. وبالتالي عدد المعلمات المقدرة في كل معادلة هو سبعة. نتائج تقديرات OLS لنموذج VAR معطاة في جدول (4.22). اختبارات F المعطاة في هذا الجدول، استخدمت لاختبار فرض أن معاملات الفترات الزمنية المتأخرة جميعًا يساوي الصفر. وبالتالي اختبار F للمتغير x (نسبة التغير في السعر الحقيقي للبترول) يوضح أن كلاً من الفترتين الزمنيتين المتأخرتين لـ x مختلفان عن الصفر إحصائيًا، احتمال الحصول على قيمة لـ F تساوي 12.5536 تحت صحة الفرض العدمي القائل بأنها آنيًا يساويان الصفر معًا صغير جداً، حوالي 0.00004. على الجانب الآخر، فإن قيم Y في الفترتين الزمنيتين المتأخرتين معًا (نسبة التغير في الوظائف غير الزراعية في تكساس) لا تختلف معنويًا عن الصفر. لتفسير x، قيمة F هي 1.36 فقط. كل إحصاءات F الأخرى يمكن تفسيرها بشكل مماثل. على أساس هذه النتائج ونتائج أخرى مقدمة في البحث الخاص بهما، Fomby و Hirschberg استنتجوا أن المقولة الشهيرة عن الاقتصاد في تكساس ليست دقيقة، حيث إنه بعد عدم الاستقرار المبدئي الناتج من تغير أسعار البترول لـ OPEC، فإن اقتصاد تكساس الآن أقل اعتماداً على التأرجحات في أسعار البترول.

(19) Theomas B. Fomby and Joseph G. Hirschberg, "Texas in Transition: Dependence on Oil and the National Economy, "Economic Review, Federal Reserve Bank of Dallas, January 1989, pp. 11–28.

جدول (4.22) نتائج التقدير للدرجة الثانية(*) من نظام VAR لتكساس :I–1974 إلى I–1988

Dependent variable: x (percentage change in real price of oil)

Variable	Lag	Coefficient	Standard error	Significance level
x	1	0.7054	0.1409	0.8305E−5
x	2	−0.3351	0.1500	0.3027E−1
y	1	−1.3525	2.7013	0.6189
y	2	3.4371	2.4344	0.1645
z	1	3.4566	2.8048	0.2239
z	2	−4.8703	2.7500	0.8304E−1
Constant	0	−0.9983E−2	0.1696E−1	0.5589

$\bar{R}^2 = 0.2982$; $Q(21) = 8.2618$ ($P = 0.9939$)

Tests for joint significance, dependent variable = x

Variable	F-statistic	Significance level
x	12.5536	0.4283E−4
y	1.3646	0.2654
z	1.5693	0.2188

Dependent variable: y (percentage change in Texas nonagricultural employment)

Variable	Lag	Coefficient	Standard error	Significance level
x	1	0.2228E−1	0.8759E−2	0.1430E−1
x	2	−0.1883E−2	0.9322E−2	0.8407
y	1	0.6462	0.1678	0.3554E−3
y	2	0.4234E−1	0.1512	0.7807
z	1	0.2655	0.1742	0.1342
z	2	−0.1715	0.1708	0.3205
Constant	0	−0.1602E−2	0.1053E−1	0.1351

$\bar{R}^2 = 0.6316$; $Q(21) = 21.5900$ ($P = 0.4234$)

Tests for joint significance, dependent variable = y

Variable	F-statistic	Significance level
x	3.6283	0.3424E−4
y	19.1440	0.8287E−6
z	1.1684	0.3197

Dependent variable: z (percentage change in nonagricultural employment in rest of United States)

Variable	Lag	Coefficient	Standard error	Significance level
x	1	−0.8330E−2	0.6849E−2	0.2299
x	2	0.3635E−2	0.7289E−2	0.6202
y	1	0.3849	0.1312	0.5170E−2
y	2	−0.4805	0.1182	0.1828E−2
z	1	0.7226	0.1362	0.3004E−5
z	2	−0.1366E−1	0.1336	0.9190
Constant	0	−0.2387E−2	0.8241E−3	0.5701E−2

$\bar{R}^2 = 0.6503$; $Q(21) = 15.6182$ ($P = 0.7907$)

Tests for joint significance, dependent variable = z

Variable	F-statistic	Significance level
x	0.7396	0.4827
y	8.2714	0.8360E−3
z	27.9609	0.1000E−7

(*) فترتان زمنيتان متأخرتان لكل متغير

المصدر: Economic review, Federal reserve bank of Dallas, January 1989, p.21

10.22 قياس عدم الثبات في السلاسل الزمنية المالية: نماذج ARCH و GARCH

MEASURING VOLATILITY IN FINANCIAL TIME SERIES: THE ARCH AND GARCH MODELS

كما لاحظنا في مقدمة هذا الفصل، فإن السلاسل الزمنية المالية مثل أسعار الأسهم، معدلات تغير العملة، معدلات التضخم وهكذا، غالبًا توجد فيها ظاهرة العنقودية المتطايرة، بمعنى فترات زمنية يظهر فيها تأرجح كبير للأسعار، ويستمر لفترة تالية، ثم يتبع ذلك فترات بها هدوء نسبي. فكما لاحظ Philip Franses.

بما أن مثل هذه البيانات (السلاسل الزمنية المالية) تعكس نتيجة التبادل بين البائعين والمشترين، على سبيل المثال، في أسهم الأسواق، فهناك العديد من الأخبار والأحداث الاقتصادية الخارجية الأخرى التي قد تؤثر على شكل السلسلة الزمنية لتحديد الأسعار. وبما أن هذه الأخبار يمكن أن تؤدي إلى تفسيرات عديدة وأيضًا أحداث اقتصادية ما مثل كارثة البترول قد تستمر لفترة زمنية، فإننا نلاحظ غالبًا المشاهدات الكبيرة الموجبة، والمشاهدات الكبيرة السالبة في السلسلة الزمنية المالية تتجه إلى التجمع في عناقيد[20].

معرفة إمكانية عدم الثبات تعتبر أمرًا مهمًا في العديد من المجالات. فعلى سبيل المثال، الكثير من الأبحاث المهمة في الاقتصاد الكلي درست تغير التضخم خلال فترة زمنية ما. فبالنسبة لبعض صانعي القرار، التضخم في حد ذاته ليس شيئًا سيئًا ولكن اختلافه يعتبر سيئًا حيث إنه يجعل التخطيط المالي أمرًا صعبًا.

نعني الفكرة سليمة بالنسبة للصادرات، الواردات وأسواق التبادل التجاري الأجنبي، حتى أن الاختلاف في معدلات تغير العملة يعني خسارة كبيرة أو مكسبًا كبيرًا. المستثمرون في أسواق الأسهم مهتمون بشكل واضح بإمكانية تطاير أسعار الأسهم، حيث إن الإمكانية العالية لعدم الثبات تعني خسارة كبيرة أو مكسبًا كبيرًا وبالتالي حالة عدم تعيين كبيرة. في الأسواق التي توجد فيها ظاهرة عدم الثبات يكون من الصعب على الشركات زيادة رأس المال في أسواق المال.

كيف يمكن أن تتم نمذجة السلاسل الزمنية المالية التي قد تظهر فيها ظاهرة عدم الثبات هذه؟ مثلاً، كيف ننمذج السلسلة الزمنية لأسعار الأسهم، معدلات تغير

(20)Philip Hans Franses, Time Series Models for Business and Economic Forecasting, Cambridge University Press, New York, 1998, p. 155.

العملة، التضخم وهكذا؟ الصفة المميزة الموجودة في كل هذه السلاسل الزمنية المالية هي أنها من حيث الشكل تمثل عملية سير عشوائية، بمعنى أنها غير ساكنة. على الجانب الآخر، فإن فروقها الأولى ساكنة بوجه عام، كما رأينا في حالة سلسلة GDP في الفصل السابق حتى ولو أن GDP لايمثل بالضبط سلسلة زمنية مالية. وبالتالي، بدلاً من عمل نموذج للمستويات المختلفة الزمنية المالية، لماذا لانقوم بعمل نموذج للفروق الأولى؟ ولكن هذه الفروق عادة ما يظهر فيها تأرجحات وفروق كبيرة، أو يظهر فيها ظاهرة عدم الثبات، مما يعني أن تباين السلسلة الزمنية المالية يختلف بشكل كبير مع الزمن. كيف يمكن أن نقوم بعمل نموذج لمثل هذا "التباين المتغير"؟ لهذا السبب وجد ما يسمى بنماذج الانحدار الذاتي المشروط باختلاف التباين (ARCH) والذي اقترحه Engle[21].

كما يظهر من الاسم، فاختلاف التباين أو عدم تساوي قيمة قد يكون له شكل بنائي في الانحدار الذاتي، حيث يظهر اختلاف التباين خلال فترات زمنية قد تكون مرتبطة ذاتيًا. لفهم ما يعنيه ذلك دعنا نستعرض المثال التالي.

مثال:

معدل تغيير العملة U.S./ U.K.: U.S./ U.K. Exchange Rate

شكل (6.22) يوضح لوغاريتم معدل التغيير الشهري للعملة .U.S إلى .U.K (دولار لجنيه استرليني) للفترة 1973 إلى 1995 لـ 276 مفردة شهريًا. كما ترى من الشكل، هناك عدد من الارتفاعات والانخفاضات في معدل تغيير العملة خلال الفترة الزمنية للعينة. لدراسة ذلك يوضح، في شكل (7.22) رسمنا التغييرات في لوغاريتم معدل تغيير العملة، لاحظ أن التغير في لوغاريتم المتغير يرمز له بالتغير النسبي، أي عند ضربه في 100 يعطي نسبة التغير. كما يمكنك ملاحظة أن التغير النسبي في معدل تغيير العملة U.S./ U.K يظهر فترات من التأرجح الكبير لبعض الوقت، وفترات من التأرجح المتوسط في أحيان أخرى، مما يعني وجود ظاهرة العنقودية المتطايرة.

الآن السؤال العملي: كيف يمكن قياس هذا التطاير إحصائيًا؟ دعنا نوضح ذلك من خلال مثال تغيير العملة الحالي.

دع معدل تغيير العملة $Y_t = \text{U.S./U.K.}$

لوغاريتم $Y_t^* = \quad Y_t$

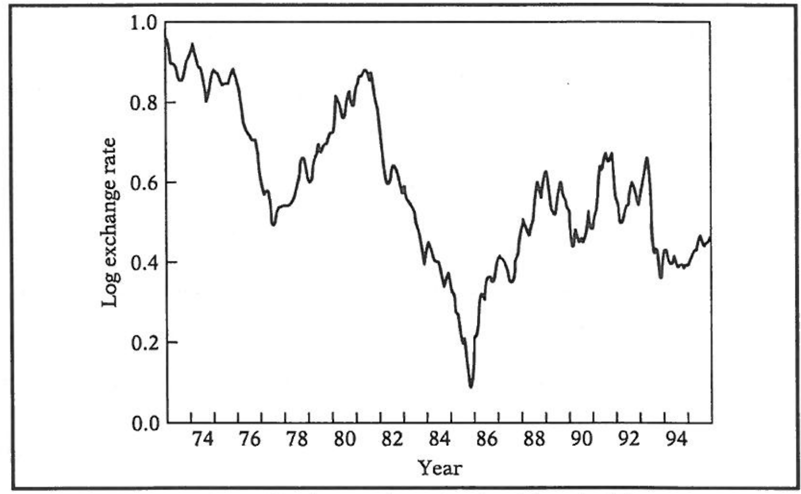
التغير النسبي في معدل تغيير العملة $dY_t^* = Y_t^* - Y_{t-1}^* =$

متوسط $d\bar{Y}_t^* = \quad dY_t^*$

$X_t = dY_t^* - d\bar{Y}_t^*$

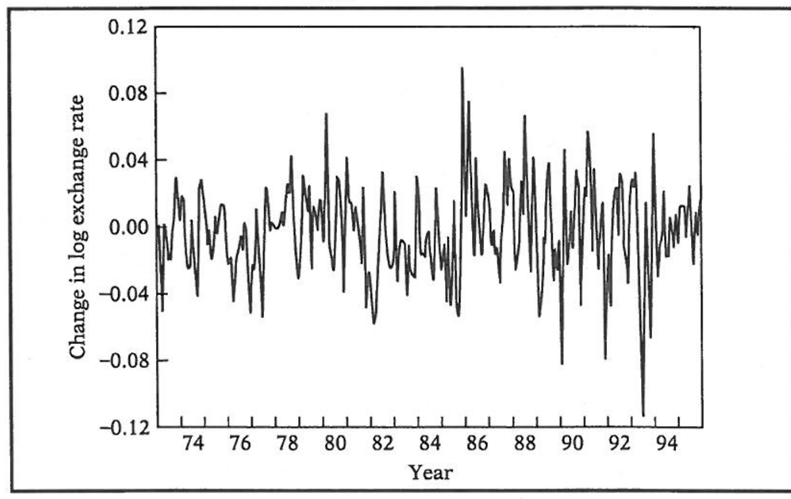
(21) R. Engle, "Autoregressive Conditional Heterosced asticity with Estimates of the Variance of United Kingdom Inflation, "Econometrica, vol. 50. no. 1, 1982, pp. 987–1007.See also A. Bera and M. Higgins, "ARCH Models: Properties, Estimation and Testing, "Journal of Economic Surveys, vol. 7, 1993, pp. 305–366.

أي أن X_t هو المتوسط المرجح في التغير النسبي لمعدل تغيير العملة. الآن نستطيع استخدام X_t^2 كمقياس لعدم الثبات. وبما أن هذه الكمية مربعة، فإن قيمتها ستكون كبيرة في الفترات التي تحدث فيها تغييرات كبيرة في أسعار الأصول المالية، وستكون قيمتها صغيرة نسبياً عندما تكون التغييرات في أسعار الأصول المالية قليلة أو متوسطة[22].



شكل (6.22) لوغاريتم معدل تغيير العملة U.S./ U.K.

1995-1973 (شهريًا)



شكل (7.22) التغير في لوغاريتم معدل تغيير العملة U.S./ U.K.

(22) قد تتساءل لماذا لم نستخدم تباين $\sum X_t^2 / n = X_t$ كمقياس للتطاير. هذا بسبب أننا نريد أن نأخذ في الاعتبار التغير في عدم ثبات أسعار الأصول مع مرور الزمن. وإذا استخدمنا تباين X_t سيكون ذلك قيمة مفردة لمجموعة معطاة في البيانات.

بعد قبول استخدام X_t^2 كمقياس لعدم الثبات، كيف يمكن معرفة أن يتغير مع مرور الزمن؟ افترض إن لدينا نموذج AR(1) أو ARIMA(1,0,0) التالي:

$$X_t^2 = \beta_0 + \beta_1 X_{t-1}^2 + u_t \qquad (1.10.22)$$

هذا النموذج يفترض أن التطاير في الفترة الزمنية الحالية مرتبط بقيمته في الفترة السابقة بالإضافة إلى مقدار خطأ عشوائي بحت. إذا كانت β_1 موجبة، فإن ذلك يعني أن التطاير كان كبيراً في الفترة السابقة، وسيظل مرتفعاً في الفترة الحالية مما يعني وجود ظاهرة العنقودية المتطايرة. إذا كان β_1 يساوي الصفر، فإنه لايوجد عنقودية في عدم الثبات. المعنوية الإحصائية للقيمة المقدرة β_1 يمكن الحكم عليها من خلال اختبار t التقليدي.

لايوجد ما يمنع من اعتبار نموذج AR(p) للتطاير كالتالي:

$$X_t^2 = \beta_0 + \beta_1 X_{t-1}^2 + \beta_2 X_{t-2}^2 + \cdots + \beta_p X_{t-p}^2 + u_t \qquad (2.10.22)$$

هذا النموذج يفترض أن عدم الثبات في الفترة الحالية مرتبط بعدم الثبات في الفترات p الزمنية السابقة. قيمة p ستعتبر سؤالاً عملياً مهماً. هذا التساؤل العملي يمكن الإجابة عنه بواحد أو أكثرطرق اختيار النماذج التي ناقشناها في الفصل (13) (مثل مقياس المعلومات لـ Akaike). يمكننا اختبار معنوية كل معامل β على حدة باستخدام اختبار t ومعنوية اثنين أو أكثر من المعاملات باستخدام اختبار F التقليدي.

نموذج (1.10.22) هو مثال لنموذج ARCH(1) و (2.10.22) يسمى نموذج ARCH(p) حيث p تمثل عدد مقادير الانحدار الذاتي في النموذج.

قبل الانتقال إلى نقطة أخرى، دعنا نشرح معدل ARCH لبيانات معدل تغيير العملة .U.S./ U.K نتائج نموذج ARCH(1) هي:

$$X_t^2 = 0.0006 + 0.1694 X_{t-1}^2$$
$$t = (6.7831) \quad (2.8355) \qquad R^2 = 0.0287 \qquad d = 1.9972 \qquad (3.10.22)$$

حيث إن X_t^2 معرفة كما سبق.

بما أن معامل مقدار الفترة الزمنية المتأخرة له معنوية عالية (P value حوالي 0.005) فهذا يعني وجود العنقودية المتطايرة في الوقت الحالي. قد استخدمنا نماذج ARCH من درجات أعلى، ولكن نموذج AR(1) كان هو النموذج الوحيد المعنوي.

كيف يمكننا اختبار تأثير ARCH في نموذج الانحدار المعتمد على بيانات سلاسل زمنية بوجه عام؟ لكي نكون أكثر تحديداً، دعنا نستعرض نموذج الانحدار الخطي الذي يحتوي على k متغير كالتالي:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t \qquad (4.10.22)$$

دعنا نفترض أنه بمعلومية البيانات في الزمن $(t - 1)$، فإن مقدار الخطأ الموزع كالتالي:

$$u_t \sim N\left[0, \left(\alpha_0 + \alpha_1 u_{t-1}^2\right)\right] \qquad (5.10.22)$$

أي أن له التوزيع الطبيعي بتوقع يساوي الصفر

$$\text{var}(u_t) = \left(\alpha_0 + \alpha_1 u_{t-1}^2\right) \qquad (6.10.22)$$

أي تباين u_t هو عملية من النوع ARCH(1).

فرض التوزيع الطبيعي لـ u_t لا يعتبر فرضًا جديداً بالنسبة لنا. الجديد هو أن تباين u عند الزمن t يعتمد على مربع الخطأ في الفترة $(t - 1)$ مما يعني وجود ارتباط تسلسلي[23]. بالطبع تباين الخطأ قد لايعتمد فقط على فترة زمنية متأخرة واحدة لمربع مقدار الخطأ ولكن على العديد من مربعات هذا المقدار في فترات زمنية متأخرة أخرى كالتالي:

$$\text{var}(u_t) = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_p u_{t-p}^2 \qquad (7.10.22)$$

إذا لم يكن هناك أي ارتباط ذاتي في تباين الخطأ يكون لدينا:

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_p = 0 \qquad (8.10.22)$$

وفي هذه الحالة يكون $\text{VAR}(u_t) = \alpha_0$ ولايكون لدينا تأثير ARCH.

وبما أننا لانلاحظ مباشرة α_t^2. Engle أوضح أنه باستخدام الانحدار التالي يمكننا بسهولة اختبار الفرض العدمي السابق:

$$\hat{u}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{u}_{t-1}^2 + \hat{\alpha}_2 \hat{u}_{t-2}^2 + \cdots + \hat{\alpha}_p \hat{u}_{t-p}^2 \qquad (9.10.22)$$

حيث إن $\hat{u}_t$، كالعادة ترمز إلى تباين OLS الذي حصلنا عليه من نموذج الانحدار الأصلي (4.10.22).

يمكننا اختبار الفرض العدمي H_0 باستخدام اختبار F التقليدي، أو بشكل بديل حساب nR^2 حيث R^2 هي معامل التحديد من الانحدار الإضافي (9.10.22) يمكن إثبات أن:

$$nR^2_{\text{asy}} \sim \chi_p^2 \qquad (10.10.22)$$

(23) ملاحظة فنية: تذكر أنه في النموذج الخطي التقليدي تباين ut كان مفترضًا أنه σ^2 والذي يعتبر في إطار موضوعنا الحالي تباينًا غير مشروط. إذا كانت $\alpha_1 < 1$ فإن شرط الاستقرار يمكن كتابته $\sigma^2 = \alpha_0 + \alpha_1 \sigma^2$ أي أن $\sigma^2 = \alpha_0 / (1 - \alpha_1)$. مما يوضح أن التباين غير المشروط لـ u لايعتمد على t ولكن يعتمد على معلمة α_1 لنموذج ARCH.

أي أنه في العينات كبيرة الحجم nR^2 تتبع توزيع كاي- المربع بدرجات حرية تساوي عدد مقادير الانحدار الذاتي في الانحدار الإضافي.

قبل البدء في الشرح، تأكد أنك تستطيع التفرقة بين الارتباط الذاتي لمقدار الخطأ كما شرحناه في الفصل (12) ونماذج ARCH. في نموذج ARCH يعتمد التباين (المشروط) لـ u_t على مقدار الخطأ (المربع) السابق مما يعطي الانطباع بوجود ارتباط ذاتي.

التغيرات في أسعار أسهم نيويورك:

New York Stock Exchange price changes

لشرح تأثير ARCH بشكل أوضح، شكل (8.22) يوضح التغير النسبي الشهري في مؤشر الـ NYSE (تغيير الأسهم في نيويورك) للفترة 1952–1995[24]. واضح من الرسم أن تغييرات الأسعار النسبية في مؤشر NYSE يتضح فيها ظاهرة عدم الثبات. وبالأخص التأرجح الكبير حول كارثة 1987 لأسعار الأسهم. لشرح عدم الثبات في الأسهم الموجود في الشكل، دعنا نستعرض النموذج البسيط التالي:

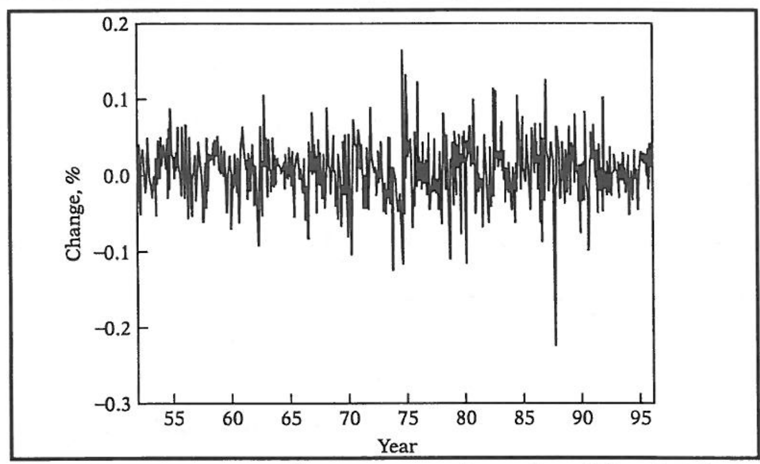
$$Y_t = \beta_1 + u_t \qquad (11.10.22)$$

حيث Y_t = التغير النسبي في مؤشر الأسهم لـ NYSE و u_t = مقدار خطأ عشوائي. لاحظ أنه بالإضافة إلى المقدار الثابت المقطوع من المحور الصادي، لا يوجد أي متغيرات مفسرة في النموذج. من البيانات حصلنا على انحدار OLS التالي:

$$\begin{aligned} \hat{Y}_t &= 0.00686 \\ t &= (3.8835) \qquad (12.10.22) \\ d &= 1.9215 \end{aligned}$$

ما الذي يرمز له هذا الجزء الثابت؟ هو ببساطة متوسط المعدل النسبي للعائد على مؤشر NYSE أو القيمة المتوسطة لـ Y_t (هل يمكنك إثبات ذلك؟). أي أنه خلال الفترة الزمنية للعينة المتوسط الشهري للعائد على مؤشر NYSE كان حوالي 0.0069%.

(24) هذا الرسم ونتائج الانحدار المعطاة معتمدة على بيانات جمعها
Gary Koop, Analysis of Economic Data, John Wiley & Sons, New York, 2000 (data from the data disk)
التغير النسبي الشهري في مؤشر أسعار الأسهم يمكن اعتباره معدل عائد المؤشر.



شكل (8.22) التغير النسبي الشهري في مؤشر أسعار NYSE، 1952-1995

الآن حصلنا على بواقي النموذج السابق، وقدرنا نموذج ARCH(1) فحصلنا على النتائج التالية:

$$\hat{u}_t^2 = 0.00145 + 0.1167\hat{u}_{t-1}^2$$

$$t = (8.8929) \quad (2.6934) \qquad (13.10.22)$$

$$R^2 = 0.0136 \qquad d = 2.0121$$

حيث إن $\hat{u}_t$ هو البواقي المقدرة من الانحدار (12.10.22)

بما أن مقدار الخطأ المربع في الفترة الزمنية المتأخرة معنوي إحصائيًا (P-value حوالي 0.007) يبدو أن تباينات الأخطاء مرتبطة، أي أن هناك تأثير ARCH. قد استخدمنا نماذج ARCH من درجات أعلى، ولكن النموذج ARCH(1) هو النموذج الوحيد المعنوي إحصائيًا.

ماذا نفعل عند وجود ARCH؛ What to do if ARCH present

تذكر أننا ناقشنا العديد من الطرق الخاصة بتصحيح اختلاف التباين، والتي تعتمد بشكل رئيسي على تطبيق OLS على البيانات المحولة. تذكر أن OLS عندما تطبق على البيانات المحولة، فإننا نسمي ذلك المربعات الصغرى العامة (GLS) إذا وجد تأثير لـ ARCH سنحتاج إلى استخدام GLS. لن نتطرق هنا إلى التفاصيل الفنية لذلك، حيث إنها تقع خارج نطاق هذا الكتاب[25]. لحسن الحظ حزم البرامج الإلكترونية مثل Eviews، Shazam، Microfit و Pc-give لديها أكواد مفهومة نوعًا ما لتقدير مثل هذه النماذج.

(25) انظر في: Russell Davidson and James G. Mackinnon, Estimation and Inference in Econometrics, Oxford University Press, New York, 1993, Sec. 16.4 and William H. Greene, Econometric Analysis, 4th ed., Prentice Hall, Englewood Cliffs, N.J., 2000, Sec. 18.5.

تعليق على إحصاء Durbin-Watson d وتأثير ARCH :
A Word on the Durbin- Watson d and the ARCH effect

سبق وأن ذكرنا القارئ في العديد من المرات، أن إحصاء d المعنوي ليس يعني بالضرورة أن هناك ارتباطًا ذاتيًا معنويًا في البيانات محل الدراسة. في العديد من الحالات قيمة d المعنوية تعتبر مؤشرًا على وجود خطأ في توصيف النموذج. وقد ناقشنا ذلك في الفصل (13). الآن لدينا خطأ في التوصيف إضافي يرجع إلى أثر ARCH. وبالتالي في انحدار السلاسل الزمنية، إذا كانت هناك قيمة معنوية للإحصاء d، فلابد أن نختبر أثر ARCH قبل أن تقبل ما تعنيه القيمة المعنوية لإحصاء d. مثال على ذلك معطى في تمرين 23.22.

ملاحظة على نموذج GARCH : A Note on GARCH model

منذ أن اكتشف في 1982، فإن نمذجة ARCH أصبحت في أهمية متزايدة مع الوضع في الاعتبار الاختلافات العديدة للنموذج الأصلي. النموذج الذي اكتسب شهرة هو نموذج الانحدار الذاتي العام المشروط باختلاف التباين (GARCH) والذي قدمه Bollerslev[26]. أبسط نماذج GARCH هو GARCH (1,1) والذي يمكن كتابته كالتالي:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \qquad (14.10.22)$$

والذي يعني أن التباين المشروط لـ u عند الزمن t يعتمد ليس فقط على مربع مقدار الخطأ في الفترة الزمنية المتأخرة السابقة (كما في ARCH(1))، ولكن أيضًا على تباينه المشروط في الفترة الزمنية المتأخرة السابقة. هذا النموذج في الحالة العامة يكتب كنموذج GARCH (p, q) حيث يوجد p فترة زمنية متأخرة لمربع مقدار الخطأ و q فترة زمنية متأخرة للتباين المشروط.

أن نتعمق في التفاصيل الفنية لهذه النماذج، حيث إنها على درجة كبيرة من التعقيد باستثناء الإشارة لنموذج GARCH (1,1) المساوي لنموذج ARCH (2)، وبالتالي فإن نموذج GARCH (p, q) يعتبر نموذجًا مماثلاً لنموذج ARCH $(p+q)$[27]. في أمثلة معدل تغيير العملة U.S./ U.K. وعائد أسهم NYSE، رأينا أن نموذج ARCH (2) لم يكن معنويًا مما يعني أن نموذج GARCH (1,1) قد لايكون مناسبًا لمثل هذه الحالات.

(26) T. Bollerslev, "Generalized Autoregressive Conditional Heteoscedasticity," Journal of Econometrics, vol. 31, 1986, pp. 307–326.

(27) لمزيد من التفاصيل، انظر Davidson and MacKinnon, op. cit., pp. 58–560.

11.22 أمثلة ختامية : CONCLUDING EXAMPLES

لختام هذا الفصل، دعنا نستعرض القليل من الأمثلة الإضافية، لتوضيح بعض النقاط التي تعرضنا لها في هذا الفصل.

العلاقة بين مؤشر طلب العمالة (HWI) ومعدل البطالة (UN) في الفترة من يناير 1969 إلى يناير 2000

The relationship between the help- wanted index (HWI) and the unemployment rate (UN) from January 1969 to January 2000

لدراسة العلاقة السببية بين HWI و UN، فهناك باحثان في شروط سوق العمل اقترحا نموذج الانحدار التالي[28]:

$$\mathrm{HWI}_t = \alpha_0 + \sum_{i=1}^{25} \alpha_i \mathrm{UN}_{t-i} + \sum_{j}^{25} \beta_j \mathrm{HWI}_{t-j} \qquad (1.11.22)$$

$$\mathrm{UN}_t = \alpha_0 + \sum_{i=1}^{25} \lambda_i \mathrm{UN}_{t-i} + \sum_{j=1}^{25} \delta_j \mathrm{HWI}_{t-j} \qquad (2.11.22)$$

لضيق المساحة، لن نستعرض هنا النتائج الفعلية لهذا الانحدار، ولكن الاستنتاج الرئيسي المستخلص من هذه الدراسة، هو أن هناك سببية مزدوجة بين هذين العاملين في سوق العمل، وهذا الاستنتاج لايتغير مع تغير طول الفترة الزمنية المتأخرة المستخدمة. بيانات AWI و UN معطاة في أسطوانة البيانات.

نموذج ARIMA لمعدل تغيير العملة ين/ الدولار يناير 1971 إلى ديسمبر 1998[29]

ARIMA modeling on the Yen/ Dollar exchange rate: January 1971 to December 1998

معدل تغيير العملة ين/ دولار (¥/ $) هو معدل تغيير عملة أساسي. من اللوغاريتم لهذا المعدل، وجد أن شكل معدل تغيير العملة هذا يتبع بالضبط شكل سلسلة زمنية غير مستقرة. ولكن باختبار الفروق الأولى، وجد أنها ساكنة. الشكل في هذا المثال يتطابق تماما مع شكل (8.22).

(28) Marc A. Giammatteo (West Point, Class of 2000), "The Relationship between the Help Wanted Index and the Unemployment Rate," unpublished term paper.

تم تعديل بعض الرموز للتناسب مع طريق الترميز في هذا الكتاب.

(29) أشكر كلاً من Gergory M. Ogborn و Marc C. Ogborn (المنطقة الغربية، فصل 2001) لتجميع وتحليل هذه البيانات.

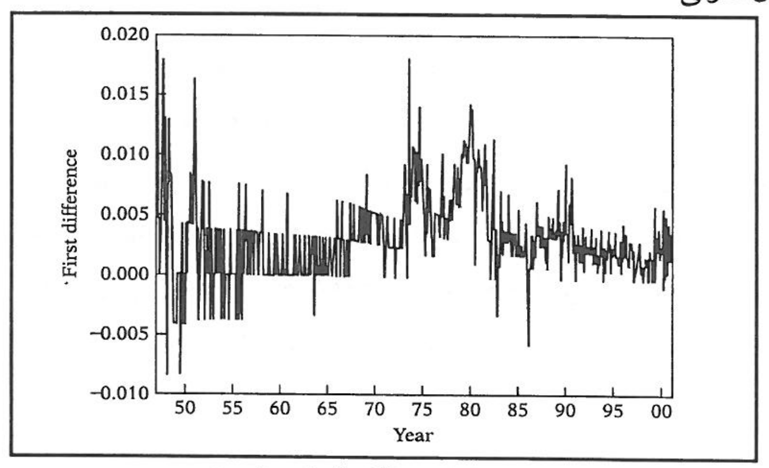
تحليل جذر الوحدة أكد أن الفروق الأولى لوغاريتم (¥/$) ساكنة. بعد اختبار مصور الارتباط لوغاريتم الفروق الأولى، قدرنا نموذج ARIMA (1,0,2) التالي:

$$\hat{Y}_t = -0.0034 + 0.9678\hat{Y}_{t-1} + -0.5866u_{t-1} - 0.4057u_{t-2}$$
$$t = (-4.3638) \quad (67.5439) \quad (-11.4361) \quad (-7.9532) \qquad (3.11.22)$$
$$R^2 = 0.1454 \qquad d = 1.9803$$

حيث إن Y_t = الفروق الأولى لوغاريتم ¥/$ و u هو مقدار الخطأ العشوائي البحت لضيق المسافة فالبيانات الخاصة بالتحليل السابق معطاة في أسطوانة البيانات. باستخدام هذه البيانات، على القارئ أن يحاول استخدام نماذج أخرى، ويقارن الأداء التنبؤي لها مع النموذج المعطى سابقًا.

نموذج ARCH لمعدل التضخم في U.S. : يناير 1947 إلى يناير 2001

لنرى ما إذا كان أثر ARCH موجود في معدل التضخم في الولايات المتحدة مقاس باستخدام CPI حيث حصلنا على بيانات CPI في الفترة من يناير 1947 إلى يناير 2001. رسم لوغاريتم الـ CPI يوضح أن السلسلة الزمنية غير ساكنة. ولكن رسم الفروق الأولى للوغاريتم الـ CPI، كما في شكل (9.22)، يوضح وجود ظاهرة التطاير حتى في الفروق الأولى الساكنة.



شكل (9.22) الفروق الأولى للوغاريتم الـ CPI

باتباع الخطوات المشروحة في انحداري (12.10.22) و (13.10.22)، قمنا أولاً بعمل انحدار للوغاريتم الفروق الأولى لـ CPI على ثابت ثم حصلنا على البواقي من هذه المعادلة، بعد تربيع هذه البواقي حصلنا على نموذج ARCH (3) التالي:

$$\widehat{\hat{u}_t^2} = 0.000052 + 0.3399\hat{u}_{t-1}^2 + 0.1338\hat{u}_{t-2}^2 + 0.0920\hat{u}_{t-3}^2$$
$$t = (5.1893) \quad (8.7270) \quad (3.5620) \quad (2.5387) \qquad (4.11.22)$$
$$R^2 = 0.2153 \qquad d = 2.0334$$

كما ترى هناك قدر من المقاومة في ظاهره التطاير، حيث إن التطاير في الشهر الحالي يعتمد على التطاير في الشهور الثلاثة السابقة. على القارئ أن يحاول الحصول على بيانات CPI من أي مصدر حكومي، ثم يحاول استخدام نموذج آخر قد يؤدي إلى نتائج أفضل، ومحتمل أن يكون هذا النموذج من نوع نماذج GARCH.

12.22 الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

1 - طرق Box- Jenkins و VAR للتنبؤ الاقتصادي، هي بدائل لنماذج المعادلة الواحدة التقليدية، ونماذج المعادلات الآنية.

2 - للتنبؤ بقيم سلسلة زمنية ما، فإن طريقة Box- Jenkins هي كالتالي:

(a) أولاً اختبر سكون السلسلة الزمنية. هذه الخطوة يمكن القيام بها بحساب دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزيئي (PACF) أو باستخدام أي تحليل تقليدي لجذر الوحدة. مصورات الارتباط الخاصة بـ ACF و PACF تعتبر دائمًا أدوات تشخيص مرئية جيدة.

(b) إذا كانت السلسلة الزمنية غير ساكنة، احصل على الفروق الأولى ذات الدرجات الأعلى للوصول إلى السكون.

(c) الـ ACF والـ PACF للسلسلة الزمنية الساكنة تحسب عادة لمعرفة ما إذا كانت السلسلة لها انحدار ذاتي فقط، أو لها متوسطات متحركة فقط، أو مزيج من الاثنين. من الخطوط المرشدة العريضة المعطاة في جدول (1.22) يمكن أن نحدد قيم p و q في عملية ARMA المناسبة. في هذه الخطوة يكون اختبار نموذج ARMA (p, q) اختبارًا تجريبيًا.

(d) يتم تقدير النموذج التجريبي.

(e) ثم اختبار بواقي هذا النموذج لمعرفة مدى عشوائيتها. إذا كانت هذه البواقي لها عشوائية بحتة، فإن النموذج التجريبي يعتبر تقريبًا جيدًا للعملية العشوائية محل الدراسة. إذا لم تكن هذه البواقي عشوائية نبدأ العملية مرة أخرى، وبالتالي فطريقة Box- Jenkins هي طريقة تكرارية.

(f) النموذج المختار في النهاية يتم استخدامه بغرض التنبؤ.

3 - طريقة VAR للتنبؤ تشتمل على عدد من السلاسل الزمنية في نفس الوقت. الخصائص المميزة لـ VAR هي كالتالي:

(a) تعتبر هذه الطريقة نظامًا آنيًا، حيث كل المتغيرات يتم التعامل معها على أنها متغيرات داخلية.

(b) في نموذج VAR قيمة المتغير يتم التعبير عنها كتوليفة خطية من قيم في الماضي أو في فترات زمنية متأخرة لهذا المتغير لكل المتغيرات الأخرى الموجودة في النموذج.

(c) إذا كانت كل معادلة في النموذج تحتوي على نفس العدد من المتغيرات في الفترات الزمنية المتأخرة، يمكن تقدير المعادلة بـ OLS بدون اللجوء إلى أي طريقة نظامية مثل المربعات الصغرى ذات المرحلتين (2SLS) أو الانحدار الذي يبدو غير مرتبط (SURE).

(d) بساطة نمذجة VAR قد تكون عيبها الرئيسي. فمع الوضع في الاعتبار العدد المحدود من المشاهدات التي تتوافر عمومًا في معظم التحاليل الاقتصادية تعتبر مسألة استخدام العديد من الفترات الزمنية المتأخرة لكل متغير من الأمور التي تستهلك العديد من درجات الحرية[30].

(e) إذا كان هناك العديد من الفترات الزمنية المتأخرة في كل معادلة، ليس من السهل دائمًا تفسير كل معامل خصوصًا إذا كانت إشارات المعاملات معكوسة. لهذا السبب يجب اختبار دالة استجابة الدافع (IRF) في نموذج VAR لمعرفة كيف يستجيب المتغير التابع للصدمات التي توجد في واحدة أو أكثر من المعادلات داخل النظام.

(f) هناك العديد من الانتقادات والشكوك حول كفاءة العديد من طرق التنبؤ. طرق مثل نماذج المعادلة المفردة، المعادلات الآنية، Box- Jenkins و VAR للتنبؤ لها مؤيدون ولها أيضًا معارضون. كل ما تستطيع قوله إنه لاتوجد طريقة واحدة مناسبة لكل المواقف. فإذا كان ذلك ممكنًا فإنه لاتوجد ضرورة

(30) المؤمنون بالإحصاء البايزي يعتقدون إمكانية تقليل حجم هذه المشكلة. انظر
R. Literman, "A Statistical Approach to Economic Forecasting, "Journal of Business and Economic Statistics, vol. 4, 1986, pp. 1–4.

لاستعراض ومناقشة كل الطرق الأخرى البديلة. وعمومًا هناك شئ مؤكد وهو أن طريقتي Box- Jenkins وVAR أصبحتا الآن من الأركان الرئيسية في علم الاقتصاد القياسي.

(4) تم في هذا الفصل دراسة فئة خاصة من هذه النماذج، ARCH، GARCH واللذان لهما أهمية خاصة في تحليل السلاسل الزمنية المالية، مثل أسعار الأسهم، معدلات التضخم، ومعدلات تغيير العملة. الشئ المميز لهذه الطرق هو أن تباين الخطأ يمكن أن يكون مرتبطًا مع بعضه البعض خلال الفترة الزمنية بسبب ظاهرة العنقودية المتطايرة. وقد أوضحنا في هذا السياق، أنه في العديد من الحالات، تكون قيمة إحصاء Durbin- Watson d المعنوية دليل على أثر ARCH أو GARCH.

تمارين: EXERCISES

أسئلة: Questions

1.22 ماهي الطرق الرئيسية للتنبؤ الاقتصادي؟

2.22 ماهي الفروق الرئيسية بين طرق المعادلات الآنية و Box- Jenkins للتنبؤ الاقتصادي؟

3.22 وضح الخطوات الخاصة بتطبيق طريقة Box- Jenkins للتنبؤ؟

4.22 ماذا يحدث إذا تم تطبيق أسلوب Box- Jenkins لسلسلة زمنية غير ساكنة؟

5.22 ماهي الفروق بين طريقة Box- Jenkins و VAR للتنبؤ الاقتصادي؟

6.22 بأي معنى يعتبر VAR غير نمطي وليس له مسار ذاتي محدد سابقًا؟

7.22 "إذا كان الهدف الأولي هو التنبؤ، VAR ستقوم بهذه المهمة". قيم هذه العبارة.

8.22 بما أن عدد الفترات الزمنية المتأخرة التي قد توجد في نموذج VAR تعتبر مسألة شخصية. كيف يمكن للفرد أن يحدد عدد الفترات الزمنية التي يجب أن يستخدمها في أي تطبيق فعلي؟

9.22 علق على العبارة التالية: "Box- Jenkins و VAR تعتبر أمثلة حقيقية للقياس بدون الاعتماد على خلفية نظرية".

10.22 ماهي العلاقة، إذا كانت هناك علاقة، بين اختبارات السببية لـ Granger ونمذجة VAR؟

مسائل: Problems

11.22 اعتبر البيانات الخاصة بـ PDI (الدخل الشخصي المتاح) المعطاة في جدول (1.21). افترض أنك تريد إيجاد نموذج ARIMA مناسبًا لهذه البيانات. حدد الخطوات اللازمة للقيام بهذه المهمة.

12.22 كرر تمرين 11.22 لبيانات الـ PCE (نفقات الاستهلاك الشخصي) المعطاة في جدول (1.21).

13.22 كرر تمرين 11.22 لبيانات الربح المعطاة في جدول (1.21).

14.22 كرر تمرين 11.22 لبيانات الأرباح المقسمة على المساهمين المعطاة في جدول (1.21).

15.22 في الفقرة 9.13 تم استعراض طريقة Schuarz لتحديد طول الفترات الزمنية المتأخرة. كيف يمكن استخدام هذه الطريقة لتحديد طول الفترات الزمنية المتأخرة المناسب في نموذج VAR؟

16.22 باستخدام بيانات PCE و PDI المعطاة في جدول (1.21) كون نموذج VAR الثنائي للفترة من I–1970 إلى IV–1990. استخدم هذا النموذج لتقدير قيم هذه المتغيرات للفترات الزمنية الربع سنوية الأربع لـ 1991، وقارن هذه القيم المتنبأ بها مع القيم الفعلية المعطاة في جدول (1.21).

17.22 كرر تمرين 16.22 باستخدام بيانات الأرباح والأرباح المقسمة على المساهمين.

18.22 باستخدام أي من حزم البرامج الإحصائية قدر دالة استجابة الدافع لفترة 8 فترات زمنية متأخرة لنموذج VAR الذي صممته في تمرين 17.22.

19.22(*) كرر تمرين 18.22 لنموذج VAR الذي صممته في تمرين 17.22.

(*) اختياري.

20.22 بالعودة إلى نتائج انحدار VAR المعطى في جدول 7.22. من قيم اختبارات F الموجودة في الانحدارات الثلاثة المعطاة، ما الذي يمكن قوله عن طبيعة السببية في المتغيرات الثلاثة.

21.22 استكمالاً لتمرين 20.22. هل يمكنك تخمين لماذا الباحثون فضلوا التعبير عن هذه المتغيرات الثلاثة الموجودة في النموذج في صورة تغييرات في النسب بدلاً من قيمها الأصلية.

22.22 باستخدام البيانات الكندية المعطاة في جدول (3.17). حدد ما إذا كان M_1 و R متغيرات عشوائية ساكنة أم لا؟ إذا لم يكونوا ساكنين. هل هناك اندماج مزدوج؟ وضح الخطوات الحسابية الأساسية.

23.22 باستخدام البيانات الموجودة في جدول (3.17). اعتبر النموذج التالي البسيط للطلب على المال في كندا:

$$\ln M_{1t} = \beta_1 + \beta_2 \ln \text{GDP}_t + \beta_3 \ln R_t + u_t$$

(a) كيف يمكنك تفسير معالم هذا النموذج؟

(b) احصل على بواقي هذا النموذج، وحدد ما إذا كان هناك أي أثر ARCH أم لا؟

24.22 بالعودة إلى نموذج ARCH(3) المعطى في (4.11.22). باستخدام نفس البيانات قدرنا نموذج ARCH(1) التالي:

$$\widehat{u_t^2} = 0.00000078 + 0.3737\hat{u}_{t-1}^2$$
$$t = (7.5843) \qquad (10.2351)$$
$$R^2 = 0.1397 \qquad d = 1.9896$$

كيف يمكنك الاختيار ما بين هذين النموذجين؟ وضح الخطوات الحسابية الأساسية؟

ملحق A

مراجعة على بعض المفاهيم الإحصائية
A REVIEW OF SOME STATISTICAL CONCEPTS

هذا الملحق ، يعتبر مقدمة تخطيطية لبعض المفاهيم الإحصائية التي تم التعرض لها في هذا الكتاب. المناقشة لن تكون دقيقة ، ولاتوجد أي إثباتات . حيث هناك العديد من الكتب الممتازة في الإحصاء ، والتي تقوم بهذه المهمة خير قيام. بعض هذه الكتب مذكورة في نهاية هذا الملحق.

1.A معاملات الجمع والضرب : SUMMATION AND PRODUCT OPERATORS

الحرف اليوناني الكبير Σ (سيجما sigma) يُستخدم للإشارة إلى الجمع ، وبالتالي فإن :

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

بعض الخصائص المهمة لمعامل الجمع Σ هي :

1- $\sum_{i=k}^{n} k = nk$ حيث k ثابت، وبالتالي $\sum_{i=1}^{4} 3 = 4 \cdot 3 = 12$.

2- $\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$ حيث k ثابت.

3- $\sum_{i=1}^{n}(a + bx_i) = na + b\sum_{i=1}^{n} x_i$ حيث a و b ثوابت، وتم استخدام الخاصية 1، 2 المذكورة أعلى.

4- $\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$.

معامل الجمع يمكن أن يتوسع ليشمل عدداً من عمليات الجمع ، وبالتالي ΣΣ أي معامل الجمع المزدوج يعرف على أنه :

$$\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}=\sum_{i=1}^{n}(x_{i1}+x_{i2}+\cdots+x_{im})$$
$$=(x_{11}+x_{21}+\cdots+x_{n1})+(x_{12}+x_{22}+\cdots+x_{n2})$$
$$+\cdots+(x_{1m}+x_{2m}+\cdots+x_{nm})$$

بعض خصائص $\Sigma\Sigma$ هي:

1 - $\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}=\sum_{j=1}^{m}\sum_{i=1}^{n}x_{ij}$ ، أي أن الترتيب في حالة الجمع المزدوج تبادلي.

2 - $\sum_{i=1}^{n}\sum_{j=1}^{m}x_i y_j=\sum_{i=1}^{n}x_i\sum_{j=1}^{m}y_j.$

3 - $\sum_{i=1}^{n}\sum_{j=1}^{m}(x_{ij}+y_{ij})=\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}+\sum_{i=1}^{n}\sum_{j=1}^{m}y_{ij}.$

4 - $\left[\sum_{i=1}^{n}x_i\right]^2=\sum_{i=1}^{n}x_i^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}x_ix_j=\sum_{i=1}^{n}x_i^2+2\sum_{i<j}x_ix_j.$

عامل الضرب Π معرَّف كالتالي:

$$\prod_{i=1}^{n}x_i=x_1\cdot x_2\cdots x_n$$

وبالتالي:

$$\prod_{i=1}^{3}x_i=x_1\cdot x_2\cdot x_3$$

2.A فراغ العينة.. نقاط العينة والأحداث:
SAMPLE SPACE, SAMPLE POINTS AND EVENTS

المجموعة التي تحتوي على كل النتائج الممكنة لتجربة عشوائية تسمى المجتمع أو فراغ العينة، وكل فرد داخل فراغ العينة يسمى نقطة العينة. وبالتالي في تجربة رمي قطعتي عملة، فراغ العينة يتكون من أربع نتائج محتملة: HH، HT، TH و TT حيث HH تعني ظهور صورة في الرمية الأولى والرمية الثانية، HT تعني ظهور صورة في الرمية الأولى وظهور كتابة في الرمية الثانية وهكذا، كل نتيجة من النتائج السابقة تسمى نقطة العينة.

الحدث هو مجموعة جزيئية من فراغ العينة. وبالتالي إذا كانت A ترمز إلى ظهور صورة وكتابة، فإنه من النتائج الممكنة السابقة يوجد اثنان فقط ينتميان إلى A وهما HT و TH. في مثل هذه الحالة، A تعتبر حدثًا، بالمثل ظهور صورتين في رمي قطعة عملة مرتين يعتبر حدثًا. الأحداث يقال عنها متنافية تبادليًا، إذا كان وقوع أحد هذه الأحداث يمنع وقوع الآخر. ففي المثال السابق ظهور HH يجعل ظهور HT في نفس

الوقت غير ممكن الحدوث. الأحداث يقال عنها مفصلة كليًا إذا اشتملت على كل النتائج الممكنة للتجربة. وبالتالي في المثال الأحداث (a) صورتان، (b) كتابتان، (c) واحدة صورة والأخرى كتابة تشتمل على كل النتائج الممكنة، وبالتالي يقال عنها أحداثًا مفصلة كليًا.

3.A الإحتمال والمتغيرات العشوائية : PROBABILITY AND RANDOM VARIABLES

الاحتمال : Probability

دع A تمثل حدثًا في فراغ العينة، $P(A)$ يرمز إلى احتمال وقوع الحدث A، أي أننا نقصد نسبة ظهور الحدث A عندما يتم تكرار التجربة. أو بمعنى آخر في مجموع n من النتائج المتساوية في إمكانية الوقوع لتجربة ما، إذا كان m منهم مفضلتين لوقوع الحدث A، فإننا نعرِّف النسبة m/n على أنها التكرار النسبي لـ A. عندما تكون n كبيرة فإن التكرار النسبي سيعطي تقريبًا جيدًا لاحتمال A.

خصائص الاحتمال : Properties of probability

$P(A)$ هو دالة حقيقية للقيمة(1) ولها الخصائص التالية:

1- لكل A $\quad 0 \le P(A) \le 1$

2 - إذا كان A، B، C، . . . تعتبر أحداثًا مفصلة كليًا، فإن $P(A + B + C + ...) = 1$، حيث $A + B + C$ تعني A أو B أو C وهكذا.

3 - إذا كانت A، B، C، . . . أحداثًا متنافية تبادليًا، فإن:

$$P(A + B + C + ...) = P(A) + P(B) + P(C) + ...$$

مثال 1

اعتبر تجربة إلقاء زهرة نرد عليها الأرقام من 1 إلى 6. فراغ العينة يتكون من 1، 2، 3، 4، . . .، 5، 6. هذه الأحداث الستة تمثل كل فراغ العينة. احتمال أي واحدة من هذه الأرقام يساوي 1/ 6 بما أنه يجد ست نتائج متساوية الإمكانية في الظهور وأي منها له احتمال متساو في الظهور. بما أن 1، 2، 3، 4، 5 و 6 تمثل كل النتائج الممكنة

(1) مدى ونطاق هذه الدالة هو مجموعة جزئية من الأعداد الحقيقية، وتسمى هذه الدالة دالة القيمة الحقيقية. لمزيد من التفاصيل انظر

Alpha C. Chiang, Fundamental Methods of Mathematical Economics, 3d ed., McGraw-Hill, 1984, Chap. 2.

للتجـربة، فإن 1 =$P(1+2+3+4+5+6)$ حيث 1، 2، 3، . . . تعني احتمال رقم 1 أو رقم 2 أو رقم 3 وهكذا. وبما أن 1، 2، . . . ، 6 تعتبر أحداثًا متنافية تبادليًا حيث لايمكن ظهور رقمين آنيًا، فإن:

$$P(1+2+3+4+5+6) = P(1)+P(2)+\ldots+P(6) = 1$$

المتغيرات العشوائية: Random Variables

المتغير الذي تتحدد قيمته بنتيجة تجربة احتمالية يسمى متغيراً عشوائيًا (rv). المتغيرات العشوائية يرمز لها عادة بالحروف الكبيرة X، Y، Z وهكذا. والقيمة التي تأخذها هذه المتغيرات يرمز لها بالحروف الصغيرة x، y، z وهكذا. المتغير العشوائي إما متقطع أو متصل. المتغير المتقطع، يأخذ قيمًا محددة (غير محددة قابلة للعد)(2). فمثلاً عند رمي زهرتي نرد كل منهما عليها الأرقام من 1 إلى 6، إذا عرفنا المتغير العشوائي X كمجموع الأرقام الظاهرة على النرد، فإن X يمكن أن تأخذ أحد هذه القيم 2، 3، 4، 5، 6، 7، 8، 9، 10، 11 أو 12. وبالتالي هذا يعتبر متغيراً متقطعًا. على الجانب الآخر، المتغير العشوائي المتصل، هو متغير يمكن أن يأخذ أي قيمة في بعض الفترات من القيم. وبالتالي طول مشاهدة ما يعتبره تغيراً متصلاً، ففي المدى مثلاً 60 إلى 65 inches يمكن أن يأخذ هذا المتغير أي قيمة معتمداً على دقة القياس.

4.A دالة كثافة الاحتمال (PDF):

PROBABILITY DENSITY FUNCTION (PDF)

دالة كثافة الاحتمال متغير عشوائي متقطع:

Probability density function of a discrete random variable

افترض أن X متغير عشوائي متقطع، يأخذ القيم X_1، X_2، . . . ، X_n، . . . فإن الدالة:

$$f(x) = P(X = x_i) \quad \text{لكل } i = 1, 2, \ldots, n, \ldots$$

$$= 0 \quad \text{لكل } x \neq x_i$$

تسمى دالة كثافة احتمال متقطعة (PDF) لـ X، حيث إن $P(X=x_i)$ يعني احتمال أن يأخذ المتغير العشوائي المتقطع X القيمة x_i.

(2) لمناقشة بسيطة حول رموز المجموعات اللانهائية القابلة للعدد انظر

R. G. D. Allen, Basic Mathematics, Macmillan, London, 1964, p. 104.

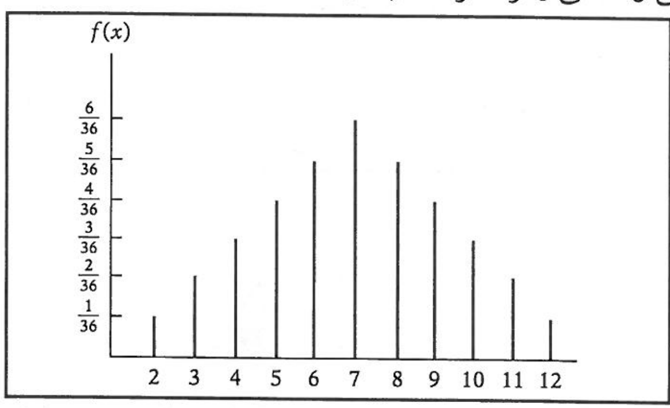
مثال 2

عند رمي زهرتي نرد، المتغير العشوائي X الذي يمثل مجموع الأرقام الظاهرة على الزهرتين، يمكن أن يأخذ قيمة من 11 قيمة الموضحة التالية. الـ pdf الخاصة بهذا المتغير موضحة كالتالي: (انظر أيضًا الشكل 1.A):

$$x = \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12$$

$$f(x) = \left(\frac{1}{36}\right)\left(\frac{2}{36}\right)\left(\frac{3}{36}\right)\left(\frac{4}{36}\right)\left(\frac{5}{36}\right)\left(\frac{6}{36}\right)\left(\frac{5}{36}\right)\left(\frac{4}{36}\right)\left(\frac{3}{36}\right)\left(\frac{2}{36}\right)\left(\frac{1}{36}\right)$$

يمكن بسهولة إثبات صحة هذه الاحتمالات. في المجموع لدينا 36 نتيجة ممكنة، حيث توجد واحدة مفضلة لرقم 2 واثنتان لرقم 3 (حيث إن المجموع 3 يمكن أن يظهر إما 1 على زهرة النرد الأولى و 2 على زهرة النرد الثانية، أو 2 على زهرة النرد الأولى و 1 على زهرة النرد الثانية) وهكذا.



شكل (1.A) دالة كثافة المتغير العشوائي المتقطع الموجود في المثال 2

دالة كثافة احتمال متغير عشوائي متصل:

Probability density function of a continuous random variable

إذا كانت X متغيراً عشوائيًا متصلاً. فإن $f(x)$ يقال عنها دالة كثافة احتمال لـ X إذا توافرت الشروط التالية:

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{a}^{b} f(x)\,dx = P(a \leq x \leq b)$$

حيث إن $f(x)\,dx$ معروفة على أنها عنصر الاحتمال (الاحتمال المرتبط بفترة صغيرة للمتغير المتصل) و $P(a \le x \le b)$ تعني احتمال وقوع X داخل الفترة a إلى b. هندسيًا معرف هذا الاحتمال كما في الشكل (2.A).

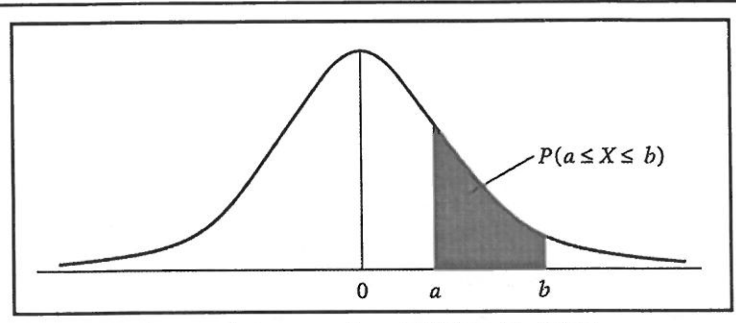
لـ rv متصل، على عكس الـ rv المتقطع، احتمال أن X يساوي قيمة محددة يساوي الصفر[3]. احتمال متغير كهذا مقاس فقط في فترة أو مدى محدد، مثل (a, b) موضح في الشكل (A.2).

مثال 3

اعتبر دالة كثافة الاحتمال التالية:

$$f(x) = \frac{1}{9}x^2 \qquad 0 \le x \le 3$$

يمكن بسهولة إثبات أن $f(x) \ge 0$ لكل قيم x داخل الفترة 0 إلى 3 وأن $\int_0^3 \frac{1}{9}\, x^2 dx = 1$ (لاحظ أن: التكامل $\frac{1}{27}x^3\Big|_0^3 = 1$). إذا أردنا حساب قيمة PDF بين مثلاً 0 و 1 نحصل على $\int_0^1 \frac{1}{9}\, x^2 dx = \left(\frac{1}{27}x^3\Big|_0^1\right) = \frac{1}{27}$ أي أن أمثال وقوع x بين 0 و 1 هو 1/27.



الشكل (2.A) كثافة احتمال متغير عشوائي متصل

دوال كثافة الاحتمال المشتركة: Joint probability density functions

PDF المشتركة المتقطة Discrete Joint PDF: دع X، Y يمثلان متغيرين عشوائيين متقطعين. وبالتالي فالدالة $f(x, y) = P(X = x, Y \ne y)$

عندما $X \ne x$ و $Y \ne y$

تعرف باسم دالة كثافة الاحتمال المشتركة المتقطعة، وهي تعطي الاحتمال (المشترك) أن تأخذ X القيمة x و Y القيمة y.

(3) لاحظ أن: $f(x)dx = 0$

مثال 4

الجدول التالي يعطي الـ PDF المشتركة للمتغيرات المتقطعة X و Y

		X			
		−2	0	2	3
Y	3	0.27	0.08	0.16	0
	6	0	0.04	0.10	0.35

هذا الجدول يعطينا احتمال أن تأخذ X القيمة 2 وفي نفس الوقت (آنيًا) تأخذ y القيمة 3 وهذا الاحتمال يساوي 0.27 واحتمال أن تأخذ X القيمة 3، بينما تأخذ Y القيمة 6 هو 0.35 وهكذا.

دالة كثافة الاحتمال الحدية: Marginal probability density function

في علاقتهم بـ $f(x, y)$ يقال إن $f(x)$ و $f(y)$ دوال كثافة احتمال حدية أو فردية، هذه الـ PDFs الحدية يتم اشتقاقها كالتالي:

$$f(x) = \sum_y f(x, y) \quad \text{PDF حدية لـ } X$$

$$f(y) = \sum_x f(x, y) \quad \text{PDF حدية لـ } Y$$

حيث إن، على سبيل المثال، Σ_y تعني التجميع على كل قيم الـ Y و Σ_x تعني التجميع على كل قيم X.

مثال 5

اعتبر البيانات المعطاة في مثال 4. PDF الحدية لـ X يتم الحصول عليها كالتالي:

$$f(x=-2) = \sum_y f(x, y) = 0.27 + 0 = 0.27$$

$$f(x=0) = \sum_y f(x, y) = 0.08 + 0.04 = 0.12$$

$$f(x=2) = \sum_y f(x, y) = 0.16 + 0.10 = 0.26$$

$$f(x=3) = \sum_y f(x, y) = 0 + 0.35 = 0.35$$

وبالمثل، دالة PDF الحدية لـ Y يتم الحصول عليها كالتالي:

$$f(y=3) = \sum_x f(x, y) = 0.27 + 0.08 + 0.16 + 0 = 0.51$$

$$f(y=6) = \sum_x f(x, y) = 0 + 0.04 + 0.10 + 0.35 = 0.49$$

كما يوضح هذا المثال للحصول على دالة PDF الحدية لـ X، نجمع الأرقام الموجودة في العمود، وللحصول على دالة PDF الحدية لـ Y نجمع الأرقام الموجودة في الصف. لاحظ أن $\Sigma_x f(x)$ لكل قيم الـ X هي 1 وأن $\Sigma_y f(y)$ لكل قيم Y هي 1 أيضًا. (لماذا؟).

PDF الشرطية Conditional PDF: كما لاحظنا في الفصل (2)، في تحليل الانحدار نكون غالبًا مهتمين بدراسة سلوك واحد من المتغيرات، مشروط بقيم المتغير أو المتغيرات الأخرى. يمكن عمل ذلك باستخدام PDF الشرطية. الدالة

$$f(x \mid y) = P(X = x \mid Y = y)$$

تسمى PDF الشرطية لـ X، حيث تعطي احتمال أن تأخذ X القيمة x، بشرط أن Y أخذت القيمة y. وبالمثل:

$$f(y \mid x) = P(Y = y \mid X = x)$$

وذلك يمثل PDF المشروطة لـ Y:

PDF المشروطة يمكن الحصول عليها كالتالي:

$$f(x \mid y) = \frac{f(x, y)}{f(y)} \quad \text{PDF المشروطة لـ } X$$

$$f(y \mid x) = \frac{f(x, y)}{f(x)} \quad \text{PDF المشروطة لـ } Y$$

وكما يتضح مما سبق، PDF المشروطة لمتغير واحد، يمكن التعبير عنها كنسبة بين PDF المشتركة إلى PDF الحدية لمتغير آخر (المشروط به).

مثال 6

باستكمال مثال 4 و 5. دعنا نحسب الاحتمال الشرطي التالي:

$$f(X = -2 \mid Y = 3) = \frac{f(X = -2, Y = 3)}{f(Y = 3)} = 0.27/0.51 = 0.53$$

لاحظ أن الاحتمال غير المشروط لـ $f(x = 2)$ هو 0.27، ولكن إذا افترضنا أن y لها القيمة المحددة 3، فإن احتمال أن X تأخذ القيمة 2- هو 0.53:

$$f(X = 2 \mid Y = 6) = \frac{f(X = 2, Y = 6)}{f(Y = 6)} = 0.10/0.49 = 0.20$$

لاحظ أيضًا أن الاحتمال غير المشروط لأن تأخذ X القيمة 2 هو 0.26 والذي يختلف عن 0.2 والذي يساوي احتمال أن تأخذ X القيمة 2 إذا افترضنا أن y تأخذ القيمة 6.

الاستقلال الإحصائي : Statistical Independence

يقال إن المتغيرين العشوائيين X و Y مستقلان إحصائيًا إذا كان :

$$f(x, y) = f(x)\, f(y)$$

أي أن الـ PDF المشتركة يمكن التعبير عنها على أنها حاصل ضرب الـ PDFs الحدية .

مثال 7

افترض أن هناك حقيبة بها ثلاث كرات عليها الأرقام 1، 2 و 3. تم سحب كرتين بطريقة عشوائية، سحب مع الإعادة من الحقيبة (بمعنى أن الكرة الأولى المسحوبة يتم إعادتها للحقيبة قبل سحب الكرة الثانية). دع X ترمز إلى الرقم الموجود على الكرة في السحبة الأولى، و Y ترمز إلى الرقم الموجود على الكرة في السحبة الثانية. الجدول التالي يوضح الـ PDF المشتركة لـ X و Y .

		X		
		1	2	3
	1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
Y	2	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
	3	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

الآن $f(X = 1, Y = 1) = \frac{1}{3}$، $f(X = 1) = \frac{1}{3}$ (تم الحصول عليه بجمع عناصر العمود الأول) و $f(Y = 1) = \frac{1}{3}$ (تم الحصول عليها بجمع عناصر الصف الأول). بما أن $f(X, Y) = f(X)\ f(Y)$ في هذا المثال، فإنه يمكن القول بأن المتغيرين مستقلان إحصائيًا. يمكن التأكد من أنه لأي توليفة أخرى من القيم لـ X و Y في الجدول المعطى أعلى، فإن الـ PDF المشتركة تساوي حاصل ضرب الـ PDFs الحدية أو الفردية .

من الممكن إثبات أن متغيرات X، Y المعطاة في مثال 4 ليست مستقلة إحصائيًا، حيث إن حاصل ضرب دالتين PDF's حديتين لن يساوي الـ PDF المشتركة .

(لاحظ أن : $f(X, Y) = f(X)\, f(Y)$ لابد أن يكون صحيحًا لكل التوليفات الممكنة لـ X و Y إذا كان المتغيران فعلاً مستقلين إحصائيًا) .

PDF المشتركة المتصلة : Continuous joint PDF

الدالة PDF $f(x, y)$ للمتغيرين المتصلين X ، Y هي كالتالي :

$$f(x, y) \geq 0$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

$$\int_{c}^{d}\int_{a}^{b} f(x, y)\, dx\, dy = P(a \leq x \leq b, c \leq y \leq d)$$

مثال 8

اعتبر الـ PDF التالية :

$$f(x, y) = 2 - x - y \qquad 0 \leq x \leq 1; 0 \leq y \leq 1$$

من الواضح أن $f(x, y) \geq 0$، بالإضافة إلى(4) :

$$\int_0^1\int_0^1 (2 - x - y)\, dx\, dy = 1$$

الـ PDF الحدية لـ X و Y يمكن الحصول عليها كالتالي :

$$f(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \qquad \text{دالة PDF الحدية لـ } X$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \qquad \text{دالة PDF الحدية لـ } Y$$

---

(4) معادلات

$$\int_0^1\left[\int_0^1 (2 - x - y)\, dx\right] dy = \int_0^1\left[\left(2x - \frac{x^2}{2} - xy\right)\Big|_0^1\right] dy$$

$$= \int_0^1\left(\frac{3}{2} - y\right) dy$$

$$= \left(\frac{3}{2}y - \frac{y^2}{2}\right)\Big|_0^1 = 1$$

لاحظ أن: الشكل $\left(\frac{3}{2}y - y^2/2\right)\Big|_0^1$ يعني أن القيمة بين الأقواس يتم حسابها عند قيمة النهاية العليا المساوية لـ 1، وعند قيمة النهاية السفلى المساوية لـ 0. وتطرح الأخيرة من الأولى للحصول على التكامل، وبالتالي في المثال السابق، فإن النهاية هي $\left(\frac{3}{2} - \frac{1}{2}\right)$ عند $y = 1$ و 0 عند $y = 0$ تعطى قيمة التكامل مساوية لـ 1.

مثال 9

دالتان PDFs الحديتان اللتان يتم الحصول عليهما من الـ PDF المشتركة المعطاة في مثال 8 هما كالتالي:

$$f(x) = \int_0^1 f(x, y)dy = \int_0^1 (2 - x - y)dy$$

$$\left(2y - xy - \frac{y^2}{2}\right)\Big|_0^1 = \frac{3}{2} - x \qquad 0 \le x \le 1$$

$$f(y) = \int_0^1 (2 - x - y)dx$$

$$\left(2x - xy - \frac{x^2}{2}\right)\Big|_0^1 = \frac{3}{2} - y \qquad 0 \le y \le 1$$

لنرى ما إذا كان المتغيران في مثال 8 مستقلين إحصائيًا أم لا. نريد أن نعرف ما إذا كان $f(x, y) = f(x) f(y)$ أم لا. بما أن $(2 - x - y) \neq (\frac{3}{2} - x)(\frac{3}{2} - y)$ فإننا نستطيع القول بأن المتغيرين غير مستقلين إحصائيًا.

5.A خصائص التوزيعات الاحتمالية :
CHARACTERISTICS OF PROBABILITY DISTRIBUTIONS

التوزيع الاحتمالي عادة ما يتم التعبير عنه في بعض من خصائصه، المعروفة باسم عزوم التوزيع. الاثنان الأكثر شهرة للاستخدام من العزوم هما المتوسط أو القيمة المتوقعة والتباين.

القيمة المتوقعة : Expected value

القيمة المتوقعة لـ X rv المتقطع، يرمز لها بالرمز $E(X)$، وتعرف كالتالي:

$$E(X) = \sum_x xf(x)$$

حيث إن Σ_x تعني التجميع على كل قيم الـ X و $f(x)$ هو الـ PDF (المتقطع) لـ X.

مثال 10

اعتبر التوزيع الاحتمالي لمجموع الرقمين الظاهرين عند رمي زهرتي نرد المعطاة في مثال 2 (انظر الشكل 1.A). بضرب قيم X المختلفة في الاحتمالات المرتبطة بالقيم وجمع حواصل الضرب لكل المفردات، نحصل على:

$$E(X) = 2\left(\tfrac{1}{36}\right) + 3\left(\tfrac{2}{26}\right) + 4\left(\tfrac{3}{36}\right) + \cdots + 12\left(\tfrac{1}{36}\right)$$
$$= 7$$

وذلك يساوي القيمة المتوسطة لمجموع الأرقام المشاهدة عند رمي زهرتي النرد مرة واحدة.

مثال 11

احسب $E(X)$ و $E(Y)$ للبيانات المعطاة في مثال 4. رأينا أن:

x	-2	0	2	3
$f(x)$	0.27	0.12	0.26	0.35

وبالتالي:

$$\begin{aligned}E(X) &= \sum_x xf(x)\\ &= (-2)(0.27) + (0)(0.12) + (2)(0.26) + (3)(0.35)\\ &= 1.03\end{aligned}$$

وبالمثل:

y	3	6
$f(y)$	0.51	0.49

$$\begin{aligned}E(Y) &= \sum_y yf(y)\\ &= (3)(0.51) + (6)(0.49)\\ &= 4.47\end{aligned}$$

القيمة المتوقعة لـ rv المتصل كالتالي:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

الفرق الوحيد بين هذه الحالة والقيمة المتوقعة لـ rv المتقطع، هو أننا بدءاً من التجميع نستخدم التكامل.

مثال 12

دعنا نوجد القيمة المتوقعة لـ PDF المتصلة المعطاة في مثال 3

$$\begin{aligned}E(X) &= \int_0^3 x\left(\frac{x^2}{9}\right)dx\\ &= \frac{1}{9}\left[\left(\frac{x^4}{4}\right)\right]_0^3\\ &= \frac{9}{4}\\ &= 2.25\end{aligned}$$

صفات القيم المتوقعة: Properties of expected values

1- القيمة المتوقعة لثابت هي نفس الثابت. وبالتالي إذا كان b ثابتًا فإن $E(b)=b$.

2 - إذا كان a، b ثوابت فإن:

$$E(aX + b) = aE(X) + b$$

ويمكن تعميم ذلك، إذا كانت X_1، X_2، . . . ، X_n هي N من المتغيرات العشوائية و a_1، a_2، . . . ، a_n، b ثوابت، فإن:

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n + b) = a_1E(X_1) + a_2E(X_2) + ... + a_NE(X_N) + b$$

3 - إذا كان X، Y متغيرين عشوائيين مستقلين، فإن:

$$E(XY) = E(X)\,E(Y)$$

بمعنى أن توقع حاصل ضرب XY يساوي حاصل ضرب توقع X (بمفردها) في توقع Y (بمفردها).

4 - إذا كان X متغيراً عشوائيًا له PDF هي $f(x)$ وإذا كان $g(X)$ أي دالة أخرى في X، فإن:

إذا كان X متقطعًا
$$E[g(X)] = \sum_x g(X)f(x)$$

إذا كان X متصلاً
$$= \int_{-\infty}^{\infty} g(X)f(x)\,dx$$

وبالتالي إذا كان $g(X) = X^2$ فإن:

إذا كان X متقطعًا
$$E(X^2) = \sum_x x^2 f(X)$$

إذا كان X متصل
$$= \int_{-\infty}^{\infty} x^2 f(X)\,dx$$

مثال 13

اعتبر الـ PDF التالية:

x	-2	1	2
$f(x)$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{2}{8}$

وبالتالي فإن:

$$E(X) = -2\left(\tfrac{5}{8}\right) + 1\left(\tfrac{1}{8}\right) + 2\left(\tfrac{2}{8}\right) = -\tfrac{5}{8}$$

و

$$E(X^2) = 4\left(\tfrac{5}{8}\right) + 1\left(\tfrac{1}{8}\right) + 4\left(\tfrac{2}{8}\right) = \tfrac{29}{8}$$

التباين: Variance

افترض أن X متغير عشوائي، و $E(X) = \mu$. التوزيع أو انتشار قيم X حول القيمة المتوقعة يتم قياسه بالتباين، والذي يعرف كالتالي:

$$\text{var}(X) = \sigma_X^2 = E(X-\mu)^2$$

الجذر التربيعي الموجب لـ σ_x^2، σ_x، يعرف على أنه الانحراف المعياري لـ X. التباين أو الانحراف المعياري يعطي مؤشرًا عن مدى قرب أو بعد مفردات الـ X عن قيمتها المتوسطة.

التباين المعرف سابقًا يحسب كالتالي:

$$\text{var}(X) = \sum_x (X-\mu)^2 f(x)$$ إذا كانت X متغيرًا عشوائيًا متقطعًا

$$= \int_{-\infty}^{\infty} (X-\mu)^2 f(x)\,dx$$ إذا كانت X متغيرًا عشوائيًا متصلاً

لتسهيل الحساب، فإن صيغة التباين المعطاة أعلى يمكن كتابتها كالتالي:

$$\begin{aligned}\text{var}(X) = \sigma_x^2 &= E(X-\mu)^2 \\ &= E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2\end{aligned}$$

بتطبيق هذه الصيغة، يمكن أن نرى أن تباين المتغير العشوائي المعطى في مثال 13 هو $\frac{29}{8} - \left(-\frac{5}{8}\right)^2 = \frac{207}{64} = 3.23$

مثال 14

دعنا نحسب تباين المتغير العشوائي المعطى في مثال 3.

$$\text{var}(X) = E(X^2) - [E(X)]^2$$

والآن:

$$\begin{aligned}E(X^2) &= \int_0^3 x^2\left(\frac{x^2}{9}\right)dx \\ &= \int_0^3 \frac{x^4}{9}dx \\ &= \frac{1}{9}\left[\frac{x^5}{5}\right]_0^3 \\ &= 243/45 \\ &= 27/5\end{aligned}$$

بما أن: $E(X) = \frac{9}{4}$ (انظر مثال 12)، لدينا في النهاية التالي:

$$\begin{aligned}\text{var}(X) &= 243/45 - \left(\frac{9}{4}\right)^2 \\ &= 243/720 = 0.34\end{aligned}$$

خصائص التباين : Properties of variance

1 - $E(X-\mu)^2 = E(X^2) - \mu^2$ ، كما لاحظنا من قبل .

2 - تباين المقدار الثابت يساوي صفراً .

3 - إذا كانت a ، b ثوابت ، فإن :

$$\text{Var}(aX+b) = a^2\ \text{var}(X)$$

4- إذا كانت X و Y متغيرين عشوائيين مستقلين فإن :

$$\text{Var}\ (X+Y) = \text{var}\ (X) + \text{var}\ (Y)$$

$$\text{Var}\ (X-Y) = \text{var}\ (X) + \text{var}\ (Y)$$

ويمكن تعميم ذلك لأكثر من متغيرين .

5 - إذا كان X و Y مستقلين و a و b ثوابت ، فإن :

$$\text{Var}\ (aX+bY) = a^2\ \text{var}\ (X) + \text{b}^2\ \text{var}\ (Y)$$

التغاير : Covariance

دع أن X و Y متغيران عشوائيان لهما القيمة المتوسطة μ_x ، μ_y بالترتيب . بالتالي فإن التغاير بين المتغيرين يتم تعريفه كالتالي :

$$\text{cov}\ (X,\ Y) = E\{(X-\mu_x)(Y-\mu_y)\} = E(XY) - \mu_x\mu_y$$

من السهل أن نرى أن تباين المتغير هو تغاير هذا المتغير مع نفسه . التغاير يتم حسابه كالتالي :

$$\text{cov}\,(X,Y) = \sum_y \sum_x (X-\mu_x)(Y-\mu_y)f(x,y)$$

$$= \sum_y \sum_x XYf(x,y) - \mu_x\mu_y$$

إذا كان X ، Y متغيرات عشوائية متقطعة و

$$\text{cov}\,(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X-\mu_x)(Y-\mu_y)f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}XYf(x,y)\,dx\,dy - \mu_x\,\mu_y$$

إذا كان X ، Y متغيرات عشوائية متصلة .

خصائص التغاير : Properties of Covariance

1 – إذا كان X، Y مستقلين، فإن تغايرهما يساوي الصفر، وبالتالي :

$$\text{cov}(X, Y) = E(XY) - \mu_x\mu_y$$

$$= \mu_x\mu_y - \mu_x\mu_y = 0$$

بما أن $E(XY) = E(X)\,E(Y) = \mu_x\mu_y$ عندما يكون X، Y مستقلين .

2 – $\text{cov}(a + bX, c + dY) = bd\,\text{cov}(X, Y)$

عندما يكون a، b، c و d ثوابت .

مثال 15

دعنا نحسب التغاير بين المتغيرين العشوائيين المتقطعين X و Y اللذين لهما الـ PDF المشتركة المعطاة في مثال 4 . من مثال 11 نحن نعرف أن :

$$\mu_y = E(Y) = 4.47 \text{ و } \mu_x = E(X) = 1.03$$

$$E(XY) = \sum_y \sum_x XYf(x, y)$$

$$= (-2)(3)(0.27) + (0)(3)(0.08) + (2)(3)(0.16) + (3)(3)(0)$$

$$= (-2)(6)(0) + (0)(6)(0.04) + (2)(6)(0.10) + (3)(6)(0.35)$$

$$= 6.84$$

وبالتالي :

$$\text{cov}(X, Y) = E(XY) - \mu_x\mu_y$$

$$= 6.84 - (1.03)(4.47)$$

$$= 2.24$$

معامل الارتباط : Correlation Coefficient

معامل ارتباط (المجتمع) ρ (rho) يعرف كالتالي :

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\{\text{var}(X)\,\text{var}(Y)\}}} = \frac{\text{cov}(X, Y)}{\sigma_x\sigma_y}$$

وبالتالي، فإن ρ يقيس الارتباط الخطي بين المتغيرين وهو يقع بين 1− و 1 .

1− تعني ارتباطًا سلبيًا تامًا و 1+ تعني ارتباطًا موجبًا تامًا .

من المعادلة السابقة يمكن أن نرى التالي :

$$\text{cov}(X, Y) = \rho\sigma_x\sigma_y$$

مثال 16

قدر معامل الارتباط للبيانات الموجودة في مثال 4.

من الـ PDFs المعطاة في مثال 11. يمكن إثبات بسهولة أن $\sigma_x = 2.05$ و $\sigma_y = 1.50$ وقد رأينا أن $\text{cov}(X, Y) = 2.24$ وبالتالي بتطبيق المعادلة السابقة، فإن

ρ يساوي $2.24\backslash(2.05)(1.5) = 0.73$.

تباين المتغيرات المرتبطة : Variances of Correlated Variables

دع أن X، Y متغيران عشوائيان، وبالتالي فإن:

$$\begin{aligned}\text{var}(X+Y) &= \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X,Y)\\ &= \text{var}(X) + \text{var}(Y) + 2\rho\sigma_x\sigma_y\\ \text{var}(X-Y) &= \text{var}(X) + \text{var}(Y) - 2\,\text{cov}(X,Y)\\ &= \text{var}(X) + \text{var}(Y) - 2\rho\sigma_x\sigma_y\end{aligned}$$

إذا كان X، Y مستقلين فإن $\text{cov}(X, Y) = 0$، وفي هذه الحالة، فإن $\text{var}(X+Y)$ و$\text{var}(X-Y)$ يتساويان معًا ويساويان $\text{var}(X) + \text{var}(Y)$ كما رأينا من قبل. النتيجة السابقة يمكن تعميمها كالتالي: إذا كانت $\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$

فإن تباين التوليفة الخطية ΣX_i هو:

$$\begin{aligned}\text{var}\left(\sum_{i=1}^{n} x_i\right) &= \sum_{i=1}^{n} \text{var}\,X_i + 2\sum_{i<j}\sum \text{cov}(X_i, X_j)\\ &= \sum_{i=1}^{n} \text{var}\,X_i + 2\sum_{i<j}\sum \rho_{ij}\sigma_i\sigma_j\end{aligned}$$

حيث إن ρ_{ij} هو معامل الارتباط بين X_i و X_j و σ_i و σ_j هما الانحراف المعياري لكل من X_i و X_j.

وبالتالي:

$$\begin{aligned}\text{var}(X_1+X_2+X_3) &= \text{var}\,X_1 + \text{var}\,X_2 + \text{var}\,X_3 + 2\,\text{cov}(X_1,X_2)\\ &\quad + 2\,\text{cov}(X_1,X_3) + 2\,\text{cov}(X_2,X_3)\\ &= \text{var}\,X_1 + \text{var}\,X_2 + \text{var}\,X_3 + 2\rho_{12}\sigma_1\sigma_2\\ &\quad + 2\rho_{13}\sigma_1\sigma_3 + 2\rho_{23}\sigma_2\sigma_3\end{aligned}$$

حيث إن σ_1 و σ_2 و σ_3 هي بالترتيب الانحراف المعياري لـ X_1، X_2 و X_3 و ρ_{12} هو معامل الارتباط بين X_1 و X_2، ρ_{13} هو معامل الارتباط بين X_1 و X_3، ρ_{23} هو معامل الارتباط بين X_2 و X_3.

التوقع الشرطي والتباين الشرطي: Conditional Expectation and Conditional Variance

افترض أن $f(x, y)$ هي الـ PDF المشتركة للمتغيرات العشوائية X و Y، التوقع الشرطي لـ X بمعلومية $Y = y$ يعرف كالتالي:

$$E(X \mid Y = y) = \sum_x x f(x \mid Y = y) \quad \text{إذا كان } X \text{ متقطعًا}$$

$$= \int_{-\infty}^{\infty} x f(x \mid Y = y)\, dx \quad \text{إذا كان } X \text{ متصلاً}$$

حيث إن $E(X \mid Y = y)$ يعني التوقع الشرطي لـ X بمعلومية أن $Y = y$ و $f(x \mid Y = y)$ هي الـ PDF الشرطية لـ X. التوقع الشرطي لـ Y، $E(Y \mid X = x)$ يتم تعريفه بطريقة مماثلة.

التوقع الشرطي Conditional Expectation: لاحظ أن $E(X \mid Y)$ هو متغير عشوائي، حيث إنه دالة في المتغير المشروط به Y عمومًا $E(X \mid Y = y)$ عندما تكون y قيمة محددة لـ Y يعتبر ثابتًا.

التباين الشرطي Conditional Variance: التباين الشرطي لـ X بمعلومية $Y = y$ يعرف كالتالي:

$$\text{var}(X \mid Y = y) = E\{[X - E(X \mid Y = y)]^2 \mid Y = y\}$$

$$= \sum_x [X - E(X \mid Y = y)]^2 f(x \mid Y = y) \quad \text{إذا كان } X \text{ متقطعًا}$$

$$= \int_{-\infty}^{\infty} [X - E(X \mid Y = y)]^2 f(x \mid Y = y)\, dx \quad \text{إذا كان } X \text{ متصلاً}$$

مثال 17

احسب $E(Y \mid X = 2)$ و $\text{Var}(Y \mid X = 2)$ للبيانات المعطاة في مثال 4

$$E(Y \mid X = 2) = \sum_y y f(Y = y \mid X = 2)$$

$$= 3f(Y = 3 \mid X = 2) + 6f(Y = 6 \mid X = 2)$$

$$= 3(0.16/0.26) + 6(0.10/0.26)$$

$$= 4.15$$

لاحظ أن:

$$f(Y=3|X=2)=f(Y=3, X=2)/f(X=2)=0.16/0.26,$$

و

$$f(Y=6|X=2)=f(Y=6, X=2)/f(X=2)=0.10/0.26,$$

وبالتالي:

$$\begin{aligned}\text{var}(Y|X=2) &= \sum_y [Y-E(Y|X=2)]^2 f(Y|X=2)\\ &= (3-4.15)^2(0.16/0.26)+(6-4.15)^2(0.10/0.26)\\ &= 2.13\end{aligned}$$

خصائص التوقع الشرطي والتباين الشرطي:

Properties of Conditional Expectation and Conditional Variance

1 - إذا كانت $f(X)$ دالة في X فإن $E(f(X)|X)=f(X)$ أي أن دالة X تتصرف كالثابت عند حساب توقعها الشرطي على X. وبالتالي فإن $[E(X^3|X]=E(X^3)$ وهكذا بسبب إذا كانت X معلومة، فإن X^3 أيضًا معلومة.

2 - إذا كانت $f(X)$ و $g(X)$ دوال في X فإن:

$$E[f(X)Y+g(X)|X]=f(X)E(Y|X)+g(X)$$

فمثلاً: $E(XY+cX^2|X)=XE(Y|X)+cX^2$ حيث c تعتبر ثابتًا.

3 - إذا كان X و Y مستقلين، فإن $E(Y|X)=E(Y)$، بمعنى أنه إذا كان X و Y متغيرين عشوائيين مستقلين، فإن التوقع الشرطي لـ Y بمعلومية X هو نفسه التوقع غير الشرطي لـ Y.

4 - قانون تكرار التوقع. من المثير للانتباه ملاحظة العلاقة التالية بين التوقع غير الشرطي للمتغير العشوائي Y، $E(Y)$ والتوقع الشرطي المعتمد على متغير آخر X، $E(Y|X)$:

$$E(Y)=E_X[E(Y|X)]$$

ويعرف هذا بقانون تكرار التوقع، مما يعني أن التوقع الحدي أو غير المشروط لـ Y يساوي توقع توقعها الشرطي، الرمز E_x يرمز إلى التوقع المأخوذ على قيم X. ببساطة، فإن هذا القانون يعني أنه إذا حصلنا أولاً على $E(Y|X)$ كدالة في X، وبعد ذلك أخذنا التوقع على توزيع قيم X، فإننا سنحصل على $E(Y)$، التوقع غير المشروط لـ Y. يمكن للقارئ إثبات ذلك باستخدام البيانات المعطاة في مثال 4.

5 -إذا كان x و y مستقلين، فإن $\mathrm{var}(Y \mid X) = \mathrm{var}(Y)$.

6 - $\mathrm{Var}(Y) = E[\mathrm{var}(Y \mid X)] + \mathrm{var}[E(Y \mid X)]$ أي أن تباين Y (غير الشرطي) يساوي توقع التباين الشرطي لـ Y بالإضافة إلى تباين التوقع الشرطي لـ Y .

العزوم الأعلى للتوزيعات الاحتمالية :

Higher moments of probability distributions

على الرغم من أن التوقع، التباين، والتغاير هي الأكثر استخداماً لمقياس مهم في حالات الـ PDFs الخاصة، بمتغير واحد أو المتعددة المتغيرات، إلا إننا أحياناً نستخدم عزوماً من درجات أعلى للـ PDFs مثل العزمين الثالث والرابع. العزمان الثالث والرابع للـ PDF المتغير الواحد $f(x)$ حول قيمة المتوسط μ يعرف كالتالي :

العزم الثالث: $E(X-\mu)^3$

العزم الرابع : $E(X-\mu)^4$

في العموم، فإن العزم r حول الوسط الحسابي يعرف كالتالي :

العزم r: $E(X-\mu)^r$

العزمان الثالث والرابع للتوزيع يستخدمان غالباً لدراسة "شكل" التوزيع الاحتمالي بالخصوص الالتواء S (بمعنى عدم التماثل)، والتفرطح (الطول أو التسطح)، وذلك موضح في الشكل (3.A).

أحد مقاييس الالتواء يعرف كالتالي :

$$S = \frac{E(X-\mu)^3}{\sigma^3}$$

$$= \frac{\text{العزم الثالث حول الوسط الحسابي}}{\text{تكعيب الانحراف المعياري}}$$

والمقياس الأكثر استخداماً للتفرطح يعرف كالتالي :

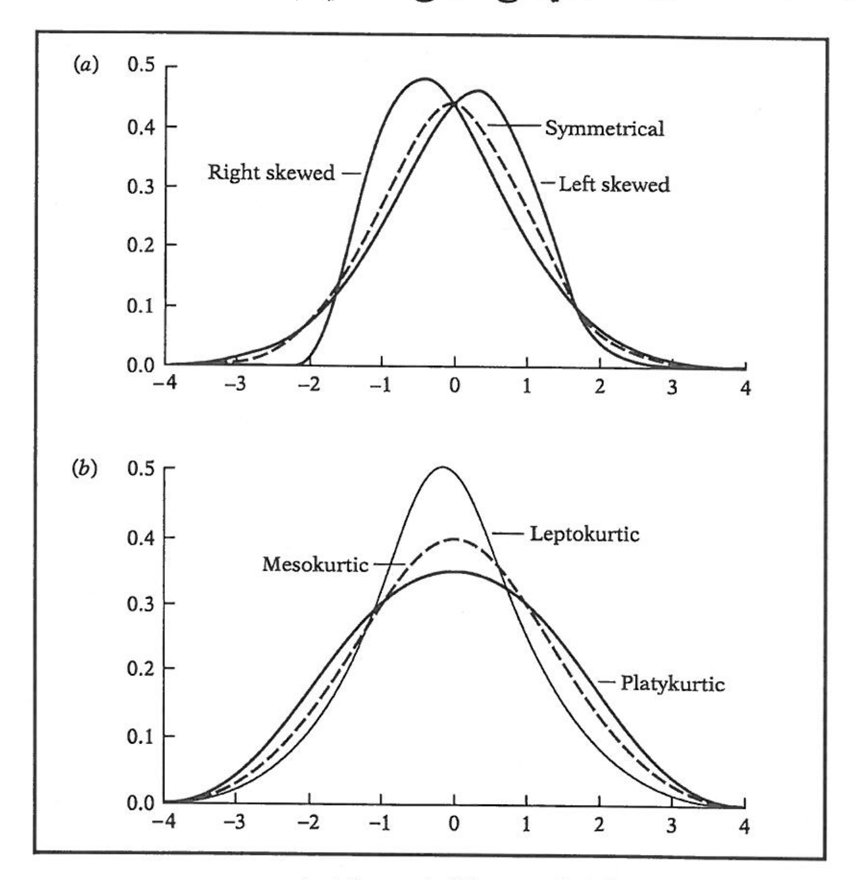
$$K = \frac{E(X-\mu)^4}{[E(X-\mu)^2]^2}$$

$$= \frac{\text{العزم الرابع حول الوسط الحسابي}}{\text{تربيع العزم الثاني}}$$

PDFs التي يكون لها قيم لـ K أقل من 3 يقال عنها platy kurtic (طرف قصير أو سميك) والأخرى التي تكون قيمة K فيها أكبر من 3 يقال عنها lepto kurtic (طرف

طويل أو رفيع). انظر الشكل (3.A) لـ PDF، يكون فيها قيمة التفرطح تساوي 3 يسمى mesokurtic ومثال عليه التوزيع الطبيعي (انظر المناقشة الخاصة بالتوزيع الطبيعي في الفقرة 6.A).

سنرى لاحقًا، كيف يمكن الدمج بين مقاييس الالتواء، ومقاييس التفرطح لتحديد ما إذا كان المتغير العشوائي يتبع التوزيع الطبيعي أم لا.



شكل (3.A) (a) الالتواء، (b) التفرطح

تذكر في عملية اختبارات الفروض، فإن اختبارات t و F معتمدة على فرض (على الأقل في العينات المحدودة أو صغيرة الحجم) إن توزيع المتغيرات محل الدراسة (أو إحصاءات العينة) هو التوزيع الطبيعي. وبالتالي يكون من الضروري جداً معرفة ما إذا كان هذا الفرض متحققًا عمليًا أم لا.

6.A بعض التوزيعات الاحتمالية النظرية المهمة : SOME IMPORTANT THEORETICAL PROBABILITY DISTRIBUTIONS

في هذا الكتاب، تم استخدام التوزيعات الاحتمالية التالية في أكثر من موضوع.

التوزيع الطبيعي : Normal distribution

التوزيع الأكثر شهرة من التوزيعات الاحتمالية النظرية، هو التوزيع الطبيعي المعروف بالشكل الجرمي، هذا الشكل المشهور حتى لدى من يمتلك معرفة إحصائية بسيطة.

المتغير العشوائي المتصل X يقال إن له التوزيع الطبيعي، إذا كانت الـ PDF الخاصة به لها الشكل التالي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \qquad -\infty < x < \infty$$

حيث μ، σ^2 معروفان باسم معالم التوزيع، وهما بالترتيب توقع وتباين التوزيع.
خصائص هذا التوزيع كالتالي:

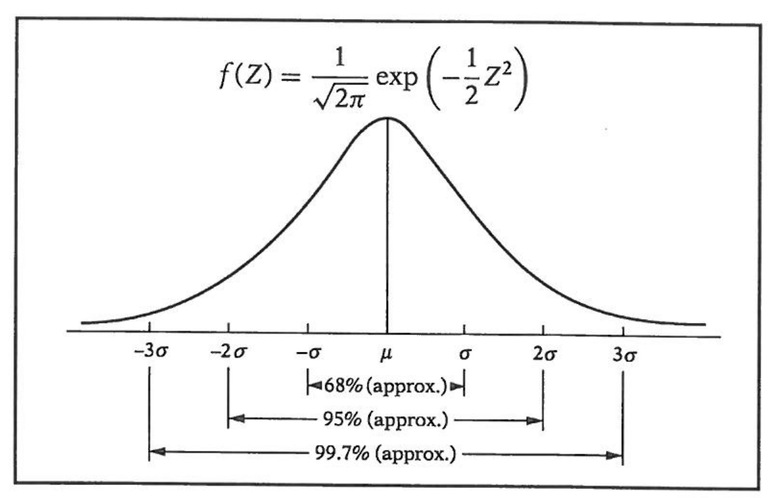
1 - توزيع متماثل حول وسطه الحسابي.

2 - 68%تقريبًا من المساحة تحت منحنى التوزيع الطبيعي تقع بين القيم $\mu\pm\sigma$ وحوالي 95% من المساحة تقع بين $\mu\pm 2\sigma$ وحوالي 99.7% من المساحة تقع بين $\mu\pm 3\sigma$ كما هو موضح في الشكل (4.A).

3 - التوزيع الطبيعي يعتمد على معلمتين μ و σ^2، وبالتالي بمجرد تحديدهما يمكن حساب احتمال أن تقع X في فترة ما باستخدام الـ PDF الخاصة بالتوزيع الطبيعي. وهذه العملية مشار إليها في جدول (1.D) في الملحق D. لاستخدام هذا الجدول نقرب المتغير الموزع توزيعًا طبيعيًا X بتوقع μ وتباين σ^2 إلى التوزيع الطبيعي القياسي Z باستخدام التحويلة التالية:

$$Z = \frac{x-\mu}{\sigma}$$

من الخصائص المهمة للمتغير القياسي، هو أن توقعه يساوي الصفر، وتباينه يساوي الوحدة. وبالتالي Z له متوسط يساوي الصفر، وتباين يساوي الواحد.
بالتعويض عن z في الـ PDF الخاصة بالتوزيع الطبيعي المعطاة سابقًا، نحصل على:



شكل (4.A) المساحات تحت المنحنى الطبيعي

وهذه الدالة تمثل الـ PDF لمتغير يتبع التوزيع الطبيعي القياسي، الاحتمالات المعطاة في ملحق D، جدول (1.D) معتمد على هذا المتغير الطبيعي القياسي.

للتسهيل نرمز للمتغير الذي له توزيع طبيعي كالتالي:

$$X \sim N(\mu , \sigma^2)$$

حيث إن ~ تعني "موزع كـ"، N ترمز إلى التوزيع الطبيعي والمقادير بين الأقواس هي معالم التوزيع الطبيعي، وهما المتوسط والتباين وبالمثل فإن:

$$X \sim N(0 , 1)$$

حيث X هي متغير طبيعي قياسي له توقع يساوي الصفر، وتباين يساوي الواحد. بعبارة أخرى، فإنه متغير طبيعي قياسي Z.

مثال 18

افترض أن $X \sim N(8,4)$. ما هو احتمال أن X تكون بين $X_1 = 4$ و $X_2 = 12$؟ لحساب الاحتمال المطلوب، نحسب قيم Z كالتالي:

$$Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{4-8}{2} = -2$$

$$Z_2 = \frac{X_2 - \mu}{\sigma} = \frac{12-8}{2} = +2$$

الآن من جدول (1.D) نلاحظ أن $\Pr(0 \le Z \le 2) = 0.4772$، ثم بالمثل لدينا $\Pr(-2 \le Z \le 0) = 0.4772$، وبالتالي فإن الاحتمال المطلوب هو $0.4772 + 0.4772 = 0.9544$ (انظر شكل 4.A).

مثال 19

ما احتمال أن الـ X في المثال السابق تزيد على 12؟

احتمال أن تزيد X على 12 هو نفس احتمال أن تزيد Z على 2. من جدول (1.D) هذا الاحتمال هو (0.4772 – 0.5) أو 0.0228.

4 – دع ($X_1 \sim N(\mu_1, \sigma_1^2)$ و ($X_2 \sim N(\mu_2, \sigma_2^2)$، وافترض أنهما مستقلان، والآن اعتبر التوليفة الخطية التالية:

$$Y = aX_1 + bX_2$$

حيث إن a، b ثوابت وبالتالي، يمكن إثبات أن:

$$Y \sim N[(a\mu_1 + b\mu_2), (a^2\sigma_1^2 + b^2\sigma_2^2)]$$

هذه النتيجة والتي تعني أن التوليفة الخطية من متغيرات طبيعية هي نفسها تتبع التوزيع الطبيعي، وذلك يمكن بسهولة تعميمه على أي توليفة خطية من اثنين أو أكثر من المتغيرات التي لها التوزيع الطبيعي.

5 – نظرية النزعة المركزية. افترض أن X_1، X_2، . . . ، X_n ترمز إلى n من المتغيرات المستقلة، كل منها له نفس الـ PDF بتوقع= μ وتباين= σ^2. دع $\bar{X} = \Sigma X_i/n$ (أي متوسط العينة)، وبالتالي كلما زاد حجم العينة n (بمعنى $n \to \infty$) فإن:

$$\underset{n\to\infty}{\bar{X}} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

أي أن $\bar{X}$ تؤول إلى التوزيع الطبيعي بتوقع u وتباين σ^2/n. لاحظ أن هذه النتيجة صحيحة بغض النظر عن شكل الـ PDF. وكنتيجة لذلك فإن:

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X} - u)}{\sigma} \sim N(0, 1)$$

وهذا هو Z متغير التوزيع الطبيعي القياسي.

6 – العزمان الثالث والرابع للتوزيع الطبيعي حول القيمة المتوسطة كالتالي:

العزم الثالث: $E(X - \mu)^3 = 0$

العزم الرابع: $E(X - \mu)^4 = 3\sigma^4$

لاحظ أن: كل العزوم الفردية حول القيمة المتوسطة للمتغيرات التي لها التوزيع الطبيعي تساوي الصفر.

7 - كنتيجة لذلك، ووفقًا لمقاييس الالتواء والتفرطح السابق ذكرها، فإن التواء PDF التوزيع الطبيعي= 0 والتفرطح= 3. بمعنى أن التوزيع الطبيعي توزيع متماثل وغير متفرطح. وبالتالي كاختبار بسيط للتوزيع الطبيعي يمكن حساب قيم الالتواء والتفرطح، ونرى إلى أي مدى هي بعيدة عن 0 و 3. هذا بالضبط هو الخلفية المنطقية لاختبار Jarque- Bera JB للتوزيع الطبيعي الذي تم مناقشته في هذا الكتاب من قبل وكان كالتالي:

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \qquad \textbf{(1.12.5)}$$

حيث S ترمز إلى الالتواء، و K التفرطح. تحت صحة الفرض العدمي، فإن JB له توزيع كاي- التربيعي X^2 بدرجات حرية 2.

8- التوقع والتباين لمتغير له التوزيع الطبيعي مستقلان، بمعنى أن أحدهما ليس دالة في الآخر.

توزيع كاي- التربيعي X^2 : The X^2 (Chi- Square) distribution

إذا كان Z_1، Z_2، . . . ، Z_k متغيرات مستقلة تتبع التوزيع الطبيعي القياسي (بمعنى إنها متغيرات تتبع التوزيع الطبيعي بتوقع يساوي الصفر، وتباين يساوي الوحدة). بالتالي فإن الكمية:

$$Z = \sum_{i=1}^{k} Z_i^2$$

يقال إنها تتبع توزيع كاي التربيعي X^2 بدرجات حرية k(df)، حيث إن المصطلح df يعني عدد الكميات المستقلة في المجموع السابق. المتغير الذي يتبع توزيع كاي التربيعي يرمز له بالرمز X_k^2 حيث إن k تساوي df. هندسيًا توزيع كاي التربيعي موضح في الشكل (5.A).

خصائص توزيع كاي التربيعي هي كالتالي:

1 - كما يتضح من الشكل (5.A)، توزيع كاي التربيعي ملتو. درجة التواء التوزيع تعتمد على df. فإذا كان هناك عدد قليل نسبيًا من df، فإن التوزيع يكون شديد الالتواء إلى ناحية اليمين، إما إذا زادت درجات الحرية فإن التوزيع يزداد تماثله. وفي حقيقة الأمر إذا زادت df عن 100 فإن المتغير:

$$\sqrt{2\chi^2} - \sqrt{(2k-1)}$$

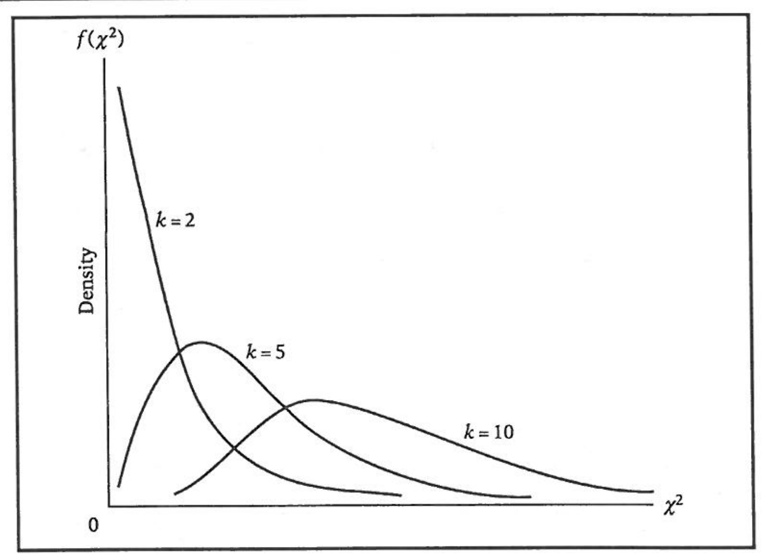
يمكن التعامل معه على أنه متغير يتبع التوزيع الطبيعي القياسي، حيث k هي درجات الحرية.

2 - القيمة المتوقعة للتوزيع كاي التربيعي هي k وتباينه $2k$، حيث k هي درجات الحرية.

3 - إذا كان Z_1 و Z_2 متغيرين مستقلين لهما توزيع كاي التربيعي بدرجات حرية k_1 و k_2، فإن المجموع $Z_1 + Z_2$ له أيضًا توزيع كاي التربيعي بدرجات حرية $\text{df} = k_1 + k_2$.

مثال 20

ما احتمال الحصول على x^2 بقيمة 40 أو أكثر بمعلومية df تساوي 20؟ من جدول (4.D) نرى أن احتمال الحصول على قيمة لـ x^2 تساوي 39.9968 أو أكثر (20 df) هي 0.005 وبالتالي احتمال الحصول على قيمة x^2 مساوية لـ 40 أو أكثر ستكون أقل من 0.005، وبالتالي هي أيضًا ذات احتمال صغير.



شكل (5.A) دالة الكثافة لمتغير يتبع توزيع كاي- التربيعي x^2

توزيع t : Student's t distribution

إذا كان Z_1 يتبع التوزيع الطبيعي القياسي [أي أن: $Z_1 \sim N(0, 1)$]، وهناك متغير آخر Z_2 يتبع توزيع كاي- التربيعي بدرجات حرية k، وموزعة مستقلاً عن Z_1، فإن المتغير المعرف كالتالي:

$$t = \frac{Z_1}{\sqrt{(Z_2/k)}}$$
$$= \frac{Z_1\sqrt{k}}{\sqrt{Z_2}}$$

يتبع توزيع t بدرجات حرية k. المتغير الذي يتبع توزيع t يرمز له عادة t_k، حيث الرمز k يعبر عن df. هندسيًا توزيع t موضح في الشكل (6.A).

خصائص توزيع t كالتالي:

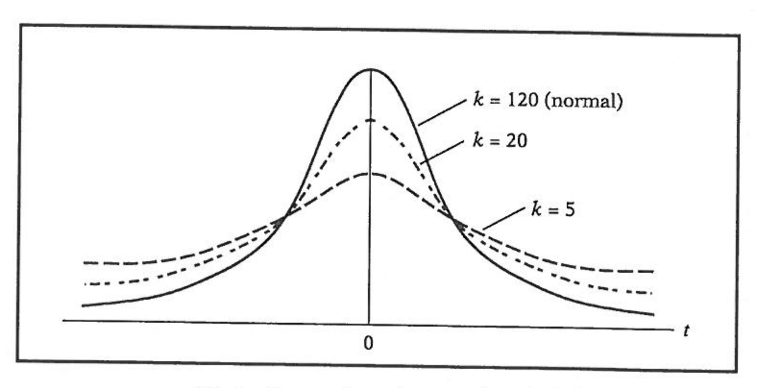
1- كما يتضح من الشكل (6.A)، فإن توزيع t، مثل التوزيع الطبيعي، توزيع متماثل، ولكنه أكثر انبساطًا من التوزيع الطبيعي. ولكن مع زيادة df، فإن توزيع t يقترب من التوزيع الطبيعي.

2- القيمة المتوقعة لتوزيع t تساوي الصفر، وتباينه يساوي $k/(k-2)$ قيم توزيع t مجدولة في جدول (2.D).

مثال 21

بافتراض أن df = 13، ما احتمال الحصول على قيمة t (a) تساوي 3 أو أكثر، (b) تساوي 3- أو أقل و (3) $|t|$ تساوي 3 أو أكثر، حيث $|t|$ تعني القيمة المطلقة لـ t (بمعنى استبعاد الإشارة)؟

من جدول (2.D)، الإجابات هي (a) حوالي 0.005، (b) حوالي 0.005 بسبب تماثل توزيع t و (c) حوالي 0.01 = 2(0.005)



شكل (6.A) توزيع t لدرجات حرية مختلفة

توزيع F : The F- distribution

إذا كان Z_1 و Z_2 مستقلين ولهما توزيع كاي التربيعي بدرجات حرية k_1 و k_2 بالترتيب فإن المتغير:

$$F = \frac{Z_1/k_1}{Z_2/k_2}$$

يتبع التوزيع F (Fisher's) بدرجات حرية k_1 و k_2. متغير توزيع F يرمز له بـ F_{k_1, k_2}، حيث إن k_1، k_2 يرمزان إلى درجات الحرية المرتبطة بمتغيرين الـ Z، حيث k_1 تسمى درجة حرية البسط و k_2 هي درجة حرية المقام. هندسيًا فإن توزيع F موضحًا في شكل (7.A).

خصائص توزيع F كالتالي:

1 - مثل توزيع كاي- التربيعي، فإن توزيع F ملتو ناحية اليمين، ولكن يمكن إثبات أنه كلما زادت k_1 و k_2 فإن توزيع F يؤول إلى التوزيع الطبيعي.

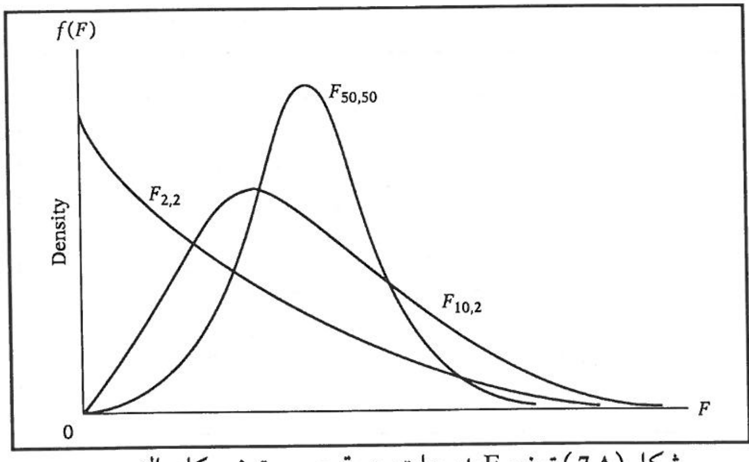
2 - القيمة المتوسطة لمتغير توزيع F هي $k_2/(k_2-2)$ والمعرفة فقط لـ $k_2 > 2$ وتباينه هو:

$$\frac{2k_2^2(k_1+k_2-2)}{k_1(k_2-2)^2(k_2-4)}$$

والمعرف لقيم $k_2 > 4$ فقط.

3 - تربيع متغير عشوائي له توزيع t بدرجات حرية k يتبع توزيع F بدرجات حرية 1 و k، بالرموز، فإن:

$$t_k^2 = F_{1,k}$$



شكل (7.A) توزيع F بدرجات حرية حسب توزيع كاي التربيعي

مثال 22

إذا كانت $k_1 = 10$ و $k_2 = 8$، ما احتمال الحصول عل قيمة F (a) تساوي 3.4 أو أكثر و (b) تساوي 5.8 أو أكثر؟

كما يتضح من جدول (3.D)، هذه الاحتمالات هي (a) تقريبًا 0.05 و (b) تقريبًا 0.01 .

4 - إذا كانت k_2، درجات حرية المقام، كبيرة نسبيًا، فإن العلاقة التالية بين توزيع F، وتوزيع كاي التربيعي تتحقق كالتالي:

$$k_1 F \sim \chi^2_{k1}$$

أي أنه بوجود درجات حرية كبيرة في المقام، فإن درجات حرية البسط مضروبة في قيمة F تقريبًا تكون مساوية لقيمة كاي- التربيعي بدرجات حرية البسط .

مثال 23

بافتراض $k_1 = 20$ و $k_2 = 100$. الـ 5% قيمة F لهذه الدرجات من الحرية هي 1.48 وبالتالي $k_1 F = (20)(1.48) = 29.6$. من توزيع كاي التربيعي بدرجات حرية 20، الـ 5% قيمة لكاي هي 31.41 .

عمومًا لاحظ أنه عند درجات الحرية الكبيرة، فإن كلاً من توزيع t، كاي- التربيعي، F يؤولان إلى التوزيع الطبيعي. هذه التوزيعات الثلاثة معروفة بالتوزيعات التي لها علاقة بالتوزيع الطبيعي.

توزيع ذو الحدين البرنولي : The Bernoulli Binomial Distribution

المتغير العشوائي X يقال إنه يتبع توزيعاً يسمى برنولي Bernoulli (عالم رياضيات سويسرى) إذا كانت دالة كثافة الاحتمال (PDF) الخاصة به هي:

$$P(X = 0) = 1 - p$$

$$P(X = 1) = p$$

حيث p، $0 \le p \le 1$، هي احتمال بعض الأحداث "النجاحات"، مثل احتمال الحصول على صورة في رمى قطعة عملة، لمثل هذا المتغير فإن:

$$E(X) = [\,1 \times p(X = 1) + 0 \times p(X = 0)] = p$$

$$\text{Var}(X) = pq$$

حيث $q = 1 - p$ وهذا يسمى احتمال الفشل.

توزيع ذو الحدين : Binomial Distribution

توزيع ذو الحدين هو الصيغة العامة لتوزيع برنولي. افترض أن n هي عدد المحاولات المستقلة، لكل منها نتائج تسمى "نجاح" فاحتمال p وأخرى "فشل" باحتمال $q = 1 - p$. إذا كانت X تمثل عدد مرات النجاح في n محاولة فإن X تتبع توزيع ذو الحدين بـ PDF كالتالي:

$$f(X) = \binom{n}{x} p^x (1-p)^{n-x}$$

حيث x تمثل عدد مرات النجاح في n محاولة، وحيث:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

حيث $n!$ تقرأ مضروب n وتعني $1 \ldots (n-2)(n-1)\, n$.

توزيع ذو الحدين هو توزيع ثنائي المعلمات، n و p. ولهذا التوزيع:

$$E(X) = np$$

$$\text{var}(X) = np\,(1-p) = npq$$

فعلى سبيل المثال، إذا رميت قطعة عملة 100 مرة، وتريد معرفة احتمال الحصول على 60 صورة، فإنك تضع $p = 0.5$، $n = 100$ و $x = 60$ في الصيغة الموجودة أعلى. هناك برامج إلكترونية في الحاسب الآلي لحساب مثل هذه الاحتمالات ترى الآن كيف أن توزيع ذو الحدين هو الحالة العامة لتوزيع برنولي.

توزيع بواسون : The Poisson Distribution

يقال إن المتغير العشوائي X له توزيع بواسون إذا كانت PDF الخاصة به كالتالي:

$$f(X) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \text{لكل } x = 0, 1, 2, \ldots, \lambda > 0$$

توزيع بواسون يعتمد على معلمة واحدة λ. والشيء المميز لتوزيع بواسون هو أن توقعه يساوي تباينه، يساوي λ، أي أن:

$$E(X) = \text{var}(X) = \lambda$$

نموذج بواسون - كما رأينا في فصل نماذج الانحدار غير الخطية - يستخدم لنمذجة الظواهر النادرة غير المتكررة الحدوث، مثل عدد المكالمات التليفونية التي

تحدث في مثلاً 5 دقائق أو عدد التذاكر التي تصدر في ساعة ما، أو عدد براءات الاختراع التي يتم الحصول عليها في مؤسسة ما خلال العام.

7.A الاستدلال الإحصائي (التقدير):
STATISTICAL INFERENCE: ESTIMATION

في فقرة 6.A درسنا العديد من التوزيعات الاحتمالية النظرية. في العديد من الحالات، نعرف أو نريد أن نفترض أن متغيراً عشوائيًا ما X يتبع احتمالي محدد، ولكن هل نعرف قيم معالم أو معلمة هذه التوزيع. فعلى سبيل المثال، إذا كانت X تتبع التوزيع الطبيعي، فنحن نريد أن نعرف قيمة معالمه الاثنين، الوسط والتباين. لتقدير المجاهيل تستخدم الطريقة المعروفة وهي افتراض أن لدينا عينة عشوائية حجمها n من التوزيع الاحتمالي المحدد ثم تستخدم بيانات العينة لتقدير المعالم المجهولة[5]. هذا معروف باسم مشكلة التقدير. في هذه الفقرة، سنقترب من هذه المشكلة أكثر، وهذه المشكلة يمكن تقسيمها إلى قسمين: التقدير بنقطة والتقدير بفترة ثقة.

التقدير بنقطة: Point Estimation

لتحديد الأفكار، دعنا نفترض أن X متغير عشوائي له PDF $f(x;\theta)$، حيث θ هي معلمة التوزيع (لتسهيل الموضوع، دعنا نفترض أن لدينا معلمة واحدة فقط مجهولة، وبسهولة يمكن تعميم ما سنتوصل إليه). افترض أننا نعلم شكل الدالة- أي نعلم الـ PDF النظرية، مثلاً توزيع t- ولكن لانعرف قيمة θ. وبالتالي نسحب عينة عشوائية من الحجم n من هذه الـPDF المعروفة، وبعد ذلك نكوِّن دالة من قيم العينة مثل:

$$\hat{\theta} = f(x_1, x_2, \ldots, x_n)$$

حتى نحصل على تقدير للمعلمة θ الحقيقية، $\hat{\theta}$ معرفة كإحصاء أو مقدر والقيمة المعينة التي يأخذها المقدر تسمى تقدير. لاحظ أن $\hat{\theta}$ يمكن التعامل معها كمتغير عشوائي، حيث إنه دالة في قيم العينة. $\hat{\theta}$ تعطينا القاعدة أن المعادلة التي تستطيع بها

(5) إذا كانت X_1، X_2، . . . ، X_n n من المتغيرات العشوائية التي لها PDF مشتركة $f(X_1, X_2, \ldots, X_n)$ إذا استطعنا كتابة

$$f(X_1, X_2, \ldots, X_n) = f(X_1)\, f(X_2) \ldots f(X_n)$$

حيث $f(X)$ هي الـPDF المشتركة لكل X فإن X_1، X_2، . . . ، X_n يقال إنها تكون عينة عشوائية من الحجم n من مجتمع له PDF هي $f(X_n)$.

تقدير θ الحقيقية. إذا كان :

$$\hat{\theta} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \bar{X}$$

حيث $\bar{X}$ هو متوسط العينة، بالتالي فإن $\bar{X}$ هو مقدر معلمة المتوسط الحقيقية، مثلاً μ، إذا كانت $\bar{X} = 50$ ، فإن ذلك يعتبر تقدير للـ μ. التقدير $\hat{\theta}$ الذي حصلنا عليه سابقًا يسمى مقدر بنقطة يعطي قيمة واحدة فقط (نقطة) كتقدير لـ θ.

التقدير بفترة : Interval Estimation

بدلاً من الحصول على نقطة تقدير واحدة لـ θ، افترض أننا حصلنا على قيمتين لـ θ باستخدام مقدرين $\hat{\theta}_2(X_1, X_2, ..., X_n)$ و $\hat{\theta}_1(X_1, X_2, ..., X_n)$ ودعنا نقول إنه بدرجة ثقة (أي احتمال) أن الفترة من $\hat{\theta}_1$ إلى $\hat{\theta}_2$ تحتوي على θ الحقيقية. وبالتالي في التقدير بفترة، على عكس التقدير بنقطة، فإننا نحصل على مدى من القيم الممكنة يحتوي على المعلمة الحقيقية θ بداخله.

الفكرة الرئيسية وراء التقدير بفترة هو استخدام العينة ، أو التوزيع الاحتمالي للمقدر. فعلى سبيل المثال، يمكن إثبات أنه إذا كان المتغير العشوائي X يتبع التوزيع الطبيعي، فإن متوسط العينة هو أيضًا يتبع التوزيع الطبيعي بتوقع= μ (التوقع الحقيقي)، وتباين= σ^2/n حيث n هو حجم العينة. بمعنى آخر التوزيع الاحتمالي للمقدر $\bar{X}$ هو $\bar{X} \sim N(\mu, \sigma^2/n)$ وبالتالي كنتيجة لذلك فإننا نكون الفترة التالية

$$\bar{X} \pm 2\frac{\sigma}{\sqrt{n}}$$

ونقول إن الاحتمال هو تقريبًا 0.95 أو 95%، إن فترات مثل السابقة ستحتوي على قيمة μ الحقيقية، حيث إننا في الحقيقة نكون فترة تقدير لـ μ. لاحظ أن الفترة المعطاة أعلى هي عشوائية، حيث إنها تعتمد على $\bar{X}$ ، والذي يختلف من عينة إلى أخرى. بوجه عام، في التقدير بفترة نكون مقدرين $\hat{\theta}_2$ و $\hat{\theta}_1$. كل منهما دالة في قيمة العينة X، مثل :

$$\Pr(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) = 1 - \alpha \qquad 0 < \alpha < 1$$

أي أنه يمكن القول بأنه باحتمال $1-\alpha$ الفترة من $\hat{\theta}_1$ إلى $\hat{\theta}_2$ تحتوي على معلمة θ الحقيقية. هذه الفترة معروفة باسم فترة ثقة من الحجم $1-\alpha$ لـ θ. $1-\alpha$ معروفة باسم معامل الثقة. إذا كانت $\alpha = 0.05$ فإن $1-\alpha = 0.95$ مما يعني أننا إذا كونا فترة ثقة بمعامل

ثقة 95% إذن بتكرار هذا التكوين من عينات أخرى متكررة، سنحصل على نتائج سليمة 95 مرة من 100 مرة، نحاول فيها أن تحتوي فترة الثقة على المعلمة الحقيقية θ. عندما يكون معامل الثقة 0.95 فنحن غالبًا نقول إن لدينا 95% فترة ثقة. عمومًا إذا كان معامل الثقة هو α −1، فإننا نقول إن لدينا $(1-\alpha)$ 100% فترة ثقة. لاحظ أن α معروفة باسم مستوى المعنوية أو احتمال حدوث الخطأ من النوع I. هذه النقطة ستتم مناقشتها في الفقرة 8.A.

مثال 24

افترض أن توزيع أطوال الرجال في مجتمع ما هو توزيع طبيعي له توقع= μ inches و σ= 2.5 inches عينة من 100 رجل تم سحبها عشوائيًا من هذا المجتمع وكان متوسط الطول 67 inches كون 95% فترة ثقة لمتوسط الطول (μ=) في المجتمع ككل.

كما سبق وذكرنا أن $\bar{X} \sim N(\mu, \sigma^2/n)$ والتي تصبح في هذه الحالة $\bar{X} \sim N(\mu, 1\ 2.5^2\backslash 100)$ من جدول (1.D) يمكن أن نستخدم التالي:

$$\bar{X} - 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

هذا يغطي 95% من المساحة تحت منحنى التوزيع الطبيعي. وبالتالي الفترة السابقة تعطي 95% فترة ثقة لـ μ. وبالتعويض بقيم $\bar{X}$، σ و n نحصل على 95% فترة ثقة كالتالي:

$$66.51 \le \mu \le 67.49$$

بتكرار مثل هذه القياسات، فترات الثقة المكونة ستحتوي على μ الحقيقية بـ 95% ثقة. النقطة الفنية التي يجب ملاحظتها هنا. هي أنه على الرغم من أننا يمكن أن نقول إن احتمال الفترة العشوائية $[\bar{X} \pm 1.96(\sigma/\sqrt{n})]$ أن تحتوي μ هو 95% لانستطيع أن نقول إن هناك احتمال 95% أن الفترة المعينة (66.49، 64.51) تحتوي على μ. بمجرد تحديد الفترة، احتمال أن تحتوي الفترة على μ إما 0 أو 1. ما نستطيع قوله هو أنه إذا كونا 100 فترة مثل هذه الفترة، فإن في 95 من الـ 100 فترة سنجد μ الحقيقية، ولكننا لانستطيع ضمان أن فترة ثقة ما بعينها ستحتوي بالضرورة على μ.

طرق التقدير: Methods of Estimation

بوجه عام، هناك ثلاث طرق لتقدير المعلمة المجهولة: (1) المربعات العنصري (LS)، (2) الإمكان الأعظم و (3) طريقة العزوم (MOM) وتعميمها طريقة العزوم العامة (GMM). قد استعرضنا طريقة LS في هذا الكتاب بشكل جيد. في الفصل (4) استعرضنا طريقة ML في إطار موضوع الانحدار، وإن كنا استخدمناها بشكل تطبيقي أكثر.

الفكرة الرئيسية وراء ML هي دالة الإمكان. لشرح ذلك، افترض أن المتغير العشوائي X له PDF $f(X, \theta)$ تعتمد على معلمة واحدة θ. نعرف الـ PDF (مثلاً برنولي أو ذو الحدين) ولكن لانعرف قيمة المعلمة المجهولة. افترض أننا حصلنا على عينة عشوائية من nX قيمة. الـ PDF المشتركة للـ n قيمة هي:

$$g(x_1, x_2, \dots, x_n; \theta)$$

حيث إنها عينة عشوائية تستطيع كتابة الـ PDF المشتركة السابقة كحاصل ضرب الـ PDF الفردية كالتالي:

$$g(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta)\ f(x_2; \theta) \dots\ f(x_n; \theta)$$

الـ PDF المشتركة لها تفسيران. إذا كانت θ معلومة، فإننا نفسرها كالاحتمال المشترك للحصول على قيم العينة المعطاة. على الجانب الآخر، يمكن التعامل معها كدالة في θ للقيم المعطاة x_1، x_2، . . . ، x_n وفي هذا التفسير الأخير نسمى الـ PDF المشتركة بدالة الإمكان (LF) ونكتبها كالتالي:

$$L(\theta; x_1, x_2, \dots, x_n) = f(x_1; \theta)\ f(x_2; \theta) \dots\ f(x_n; \theta)$$

لاحظ مكان ودور θ في دالة كثافة الاحتمال المشتركة ودالة الإمكان.

مقدر ML لـ θ هو قيمة θ التي تعظم دالة إمكان (العينة)، L.

للتسهيل الرياضي، عادة ما نأخذ لوغاريتم هذه الدالة، وتسمى لوغاريتم دالة الإمكان ($\text{Log}\ L$). باتباع قواعد الحساب للتعظيم، فإننا نفاضل دالة الإمكان بالنسبة للمعلمة المجهولة، ونساوي نتيجة التفاضل مع الصفر القيمة الناتجة للمقدر تسمى تقدير الإمكان الأعظم. من الممكن تطبيق شرط التفاضل للدرجة الثانية للحصول على القيمة العظمى للتأكد من أن النقطة التي حصلنا عليها هي القيمة العظمى للدالة.

إذا كان هناك أكثر من معلمة مجهولة، فإننا نفاضل لوغاريتم دالة الإمكان بالنسبة لكل معلمة مجهولة، ونساوي النتائج التي نحصل عليها مع الصفر، ونحل هذه المعادلات آنيًا، للحصول على قيم المعلمات المجهولة. قد أوضحنا ذلك من قبل في نموذج الانحدار المتعدد (انظر ملحق الفصل 4).

مثال 25

افترض أن المتغير العشوائي X يتبع توزيع بواسون بقيمة متوقعة λ، افترض أن X_1، X_2، . . . ، X_n متغيرات عشوائية مستقلة تتبع توزيع بواسون بمعلمة λ. افترض أننا نريد معرفة مقدر ML لـ λ. دالة الإمكان هي:

$$L(x_1, x_2, \ldots, x_n; \lambda) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \frac{e^{-\lambda}\lambda^{x_2}}{x_2!} \cdots \frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$
$$= \frac{e^{-n\lambda}\lambda^{\Sigma x_i}}{x_1!x_2!\cdots x_n!}$$

هذا التغيير غير شائع، ولكن إذا أخذنا لوغاريتم الدالة نحصل على:

$$\log(x_1, x_2, \ldots, x_n; \lambda) = -n\lambda + \sum x_i \log\lambda - \log c$$

حيث $\text{Log } c = \prod x_i!$، فاضل المعادلة السابقة بالنسبة لـ λ سنحصل على $(-n + (\Sigma x_i)/\lambda)$ وبمساواة ذلك بالصفر، نحصل على $\lambda_{ml} = (\Sigma x_i)/n = \bar{X}$ وهذا هو مقدر ML للمعلمة λ المجهولة.

طريقة العزوم The Method of moments: قد أعطينا من قبل نبذة عن MOM في تمرين 4.3 فيما يسمى مبدأ التماثل، حيث إن عزوم العينة نحاول أن تطابق خصائصها مع نظيرها من المجتمع الـ GMM، الذي يعتبر تعميمًا لـ MOM، أصبح الآن أكثر شهرة في الاستخدام، ولكن ليس في المستوى الأولي. وبالتالي لن نتطرق إليه هنا. الخصائص الإحصائية المرغوب فيها تقع في فئتين: خصائص العينات صغيرة الحجم أو المحدودة، وخصائص العينات كبيرة الحجم أو التقاربية. في هاتين الفئتين لابد من معرفة التوزيع الاحتمالي للمقدر.

خصائص العينات صغيرة الحجم Small- sample properties: عدم التحيز. المقدر $\hat{\theta}$ يقال عنه إنه مقدر غير متحيز لـ θ إذا كانت القيمة المتوقعة لـ $\hat{\theta}$ تساوي القيمة الحقيقية لـ θ، أي أن:

$$E(\hat{\theta}) = \theta$$

أو

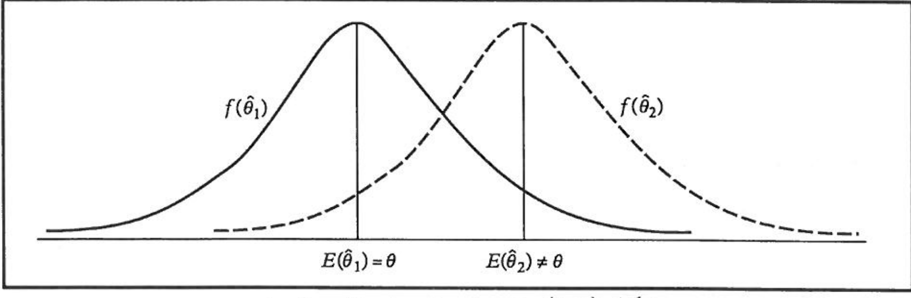
$$E(\hat{\theta}) - \theta = 0$$

إذا لم يتحقق هذا التساوي، فإن المقدار يقال عنه مقدر متحيز، ومقدر التحيز يحسب كالتالي:

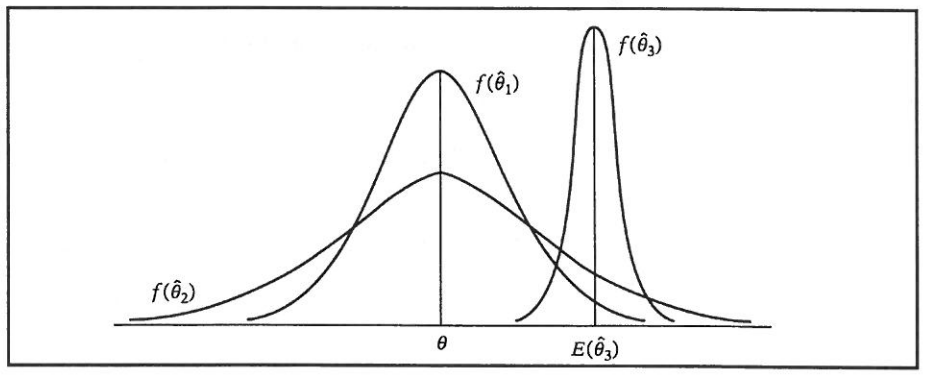
$$\text{bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$$

بالطبع إذا كان $E(\hat{\theta})= \theta$ أي أن $\hat{\theta}$ مقدر غير متحيز، فإن مقدر التحيز سيساوي الصفر.

هندسيًا، فإن ذلك معروض في الشكل (8.A). لاحظ أن عدم التحيز خاصية للعينات المتكررة، وليس بالنسبة لعينة ما محددة، مع الوضع في الاعتبار أن حجم العينة ثابت، وأننا نسحب عينات متعددة في كل مرة نحصل على مقدر المعلمة المجهولة. القيمة المتوسطة لهذه المقدرات من المتوقع أن تتساوى مع القيمة الحقيقية إذا كان المقدر غير متحيز.



شكل (8.A) مقدر غير متحيز ومقدر متحيز



شكل (9.A) توزيع ثلاثة مقدرات لـ θ

التباين الأقل: يقال إن $\hat{\theta}$ مقدر له أقل تباين إذا كان تباين $\hat{\theta}_1$ أقل من أو على الأقل يساوي تباين $\hat{\theta}_2$ والذي يعتبر مقدراً آخر لـ θ. هندسيًا في الشكل (9.A) نستعرض ثلاثة مقدرات لـ θ هي $\hat{\theta}_1$، $\hat{\theta}_2$، $\hat{\theta}_3$ وتوزيعاتها الاحتمالية. كما هو موضح تباين $\hat{\theta}_3$ هو الأقل، حيث إن تباينه أقل من كل من $\hat{\theta}_1$ و $\hat{\theta}_2$. وبالتالي بافتراض وجود الثلاثة مقدرات، فإن $\hat{\theta}_3$ هو المقدر ذو التباين الأقل، ولكن لاحظ أن $\hat{\theta}_3$ مقدر متحيز (لماذا؟).

المقدر غير المتحيز الأفضل أو المقدر الكفء: إذا كان $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدرين غير متحيزين لـ θ وتباين $\hat{\theta}_1$ أصغر من أو على الأقل يساوي تباين $\hat{\theta}_2$، فإن $\hat{\theta}_1$ هو مقدر

غير متحيز ذو تباين أقل أو مقدر غير متحيز أفضل أو مقدر كفء. وبالتالي في شكل (9.A) من المقدرين غير المتحيزين $\hat{\theta}_1$ و $\hat{\theta}_2$ فإن $\hat{\theta}_1$ هو غير المتحيز الأفضل أو المقدر الكفء.

الخطية Linearity: يقال إن $\hat{\theta}$ مقدر خطي لـ θ إذا كان دالة خطية في مفردات العينة. وبالتالي فمتوسط العينة المعرف كالتالي:

$$\bar{X} = \frac{1}{n}\sum X_i = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

يعتبر مقدر خطي، حيث إنه دالة خطية في قيم الـ X.

أفضل المقدرات الخطية غير المتحيزة Best Linear Unbiased estimator (BLUE): إذا كان $\hat{\theta}$ خطي وغير متحيز وله أقل تباين داخل فئة المقدرات الخطية غير المتحيزة لـ θ، فيقال عنه المقدر الخطي الأفضل غير المتحيز أو للاختصار BLUE.

المقدر ذو متوسط مربع الأخطاء الأقل MSE

Minimum Mean- Square error (MSE) estimator

الـ MSE لمقدر $\hat{\theta}$ تعرف كالتالي:

$$\text{MSE}(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

هذا بخلاف تباين $\hat{\theta}$ الذي يعرف كالتالي:

$$\text{var}(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2$$

الفرق بين الاثنين هو أن Var $(\hat{\theta})$ يقيس التشتت في توزيع $\hat{\theta}$ حول قيمتها المتوسطة أو القيمة المتوقعة لها أما MSE $(\hat{\theta})$ يقيس التشتت حول القيمة الحقيقية للمعلمة. العلاقة بين الاثنين كالتالي:

$$\begin{aligned}\text{MSE}(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 \\ &= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 \\ &= E[\hat{\theta} - E(\hat{\theta})]^2 + E[E(\hat{\theta}) - \theta]^2 + 2E[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta] \\ &= E[\hat{\theta} - E(\hat{\theta})]^2 + E[E(\hat{\theta}) - \theta]^2\end{aligned}$$

بما أن المقدر الأخير يساوي الصفر(6):

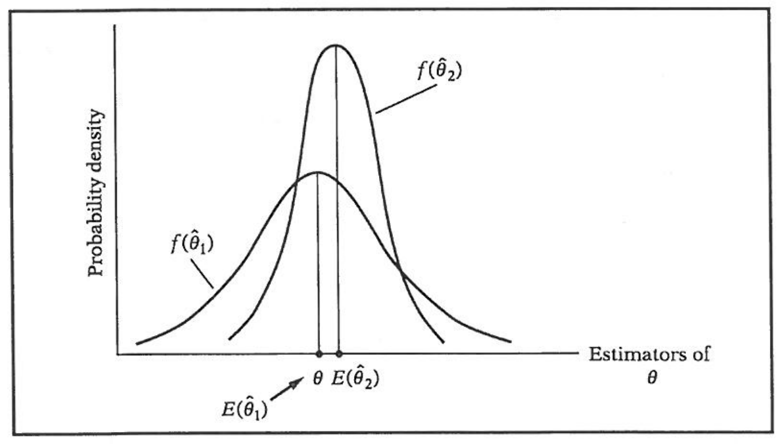
$$= \text{var}(\hat{\theta}) + \text{bias}(\hat{\theta})^2$$

(6) المقدار الأخير يمكن كتابته $2\{[E(\hat{\theta})]^2 - [E(\hat{\theta})]^2 - \theta E(\hat{\theta}) + \theta E(\hat{\theta})\} = 0$. لاحظ أيضاً أن $E[E(\hat{\theta}) - \theta]^2 = [E(\hat{\theta}) - \theta]^2$ بما أن القيمة المتوقعة للثابت هي ببساطة الثابت نفسه.

تباين$\hat{\theta}$ بالإضافة إلى مربح مقدر التحيز =

بالطبع إذا كان مقدر التحيز يساوي الصفر فإن MSE $(\hat{\theta})$= Var $(\hat{\theta})$

طريقة MSE الأقل تعتمد على اختيار المقدر الذي له MSE أقل من بين مجموع المقدرات الموجودة. ولكن لاحظ أنه إذا وجد هذا المقدر، فإن هناك ثمنًا لذلك- للحصول على التباين الأقل لابد من قبول بعض التحيز. هندسيًا فإن ذلك مشروح في شكل 10.A. في هذا الشكل، $\hat{\theta}_2$ لها تحيز بسيط، ولكن تباينها أقل من المقدر غير المتحيز $\hat{\theta}_1$. في الواقع عمومًا طريقة MSE الأقل تستخدم عندما لاتستطيع تطبيق طريقة غير المتحيز الأفضل للحصول على مقدرات لها تباين أقل.



شكل (10.A) التبادلية بين التحيز والتباين

خصائص العينات كبيرة الحجم : Large- Sample properties

في بعض الأحيان لايتوافر في المقدر واحد أو أكثر من الخصائص الإحصائية المرغوب فيها في العينات صغيرة الحجم. ولكن مع زيادة حجم العينة، فإن المقدر يقترب من هذه الخصائص الإحصائية المرغوب فيها. هذه الخصائص تسمى الخصائص التقاربية، أو خصائص العينات كبيرة الحجم عدم التحيز التقاربي. المقدر $\hat{\theta}$ يقال عنه غير متحيز تقاربيًا كمقدر لـ θ إذا كان:

$$\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$$

حيث $\hat{\theta}_n$ تعني أن المقدر محسوب على أساس عينة حجمها n و Lim تعني النهاية عندما n تؤول إلى ∞، أي أن n في تزايد مستمر. بمعنى آخر، $\hat{\theta}$ هو مقدر غير

متحيز تقاربيًا إذا كان توقعه أو قيمته المتوسطة تؤول إلى القيمة الحقيقية، مع زيادة حجم العينة. كمثال على ذلك، دعنا نستعرض المقياس التالي لتباين عينة من متغير عشوائي X:

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

يمكن إثبات أن :

$$E(S^2) = \sigma^2\left(1 - \frac{1}{n}\right)$$

حيث σ^2 هي التباين الحقيقي. من الواضح أنه في العينات صغيرة الحجم، يكون S^2 مقدر متحيز، ولكن مع زيادة n فإن $E(S^2)$ يؤول إلى σ^2 الحقيقية، وبالتالي هو مقدر غير متحيز تقاربيًا.

الاتساق $\hat{\theta}$ يقال إنه مقدر متسق إذا كان يؤول إلى القيمة الحقيقية θ، كلما زاد حجم العينة. شكل (11.A) يستعرض هذه الخاصية.

في هذا الشكل ، لدينا توزيع $\hat{\theta}$ على أساس أحجام عينات 25، 50، 80 و 100، وكما يتضح من الشكل، $\hat{\theta}$ المحسوبة على أساس $n = 25$ متحيز، حيث إن توزيع العينة الخاصة بها غير متركز حول θ الحقيقية. ولكن مع زيادة حجم العينة، فإن توزيع $\hat{\theta}$ ليس فقط يتجه إلى أن يكون أقرب تمركز حول θ (بمعنى $\hat{\theta}$ تصبح أقل تحيزا) ولكن تباينه أيضًا يصبح أقل. إذا كانت نهاية توزيع $\hat{\theta}$ (بمعنى مع زيادة n) تؤول إلى نقطة واحدة θ بمعنى أنه إذا كان توزيع $\hat{\theta}$ له تشتت أو تباين يساوي الصفر، فإننا نقول إن $\hat{\theta}$ مقدر متسق لـ θ.

بشكل علمي أكثر ، فإن $\hat{\theta}$ يقال عنه إنه مقدر متسق إذا كان احتمال القيمة المطلقة للفرق بين $\hat{\theta}$ و θ أقل من δ (مقدر صغير موجب اختياري) يؤول إلى الواحد الصحيح، رمزيًا فإن ذلك يعني:

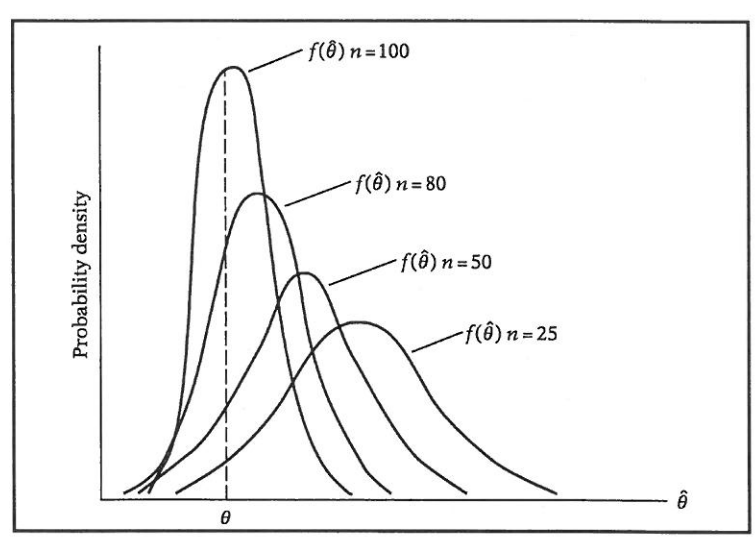
$$\lim_{n\to\infty} P\{|\hat{\theta} - \theta| < \delta\} = 1 \qquad \delta > 0$$

حيث P ترمز إلى الاحتمال. هذا يتم عادة التعبير عنه كالتالي:

$$\operatorname*{plim}_{n\to\infty} \hat{\theta} = \theta$$

حيث plim تعني نهاية الاحتمال.

لاحظ أن خاصية عدم التحيز والاتساق مختلفان تمامًا في المفهوم. خاصية عدم التحيز يمكن أن تتحقق مع أي حجم للعينة، في حين الاتساق مقصور فقط على خواص العينات كبيرة الحجم.



شكل (11.A) توزيع $\hat{\theta}$ عندما يزداد حجم العينة

الشرط الكافي للاتساق هو أن التحيز والتباين الاثنين معًا، يؤولان إلى الصفر مع زيادة حجم العينة(7). وبصياغة أخرى، فإن الشرط الكافي للاتساق هو أن MSE $(\hat{\theta})$ يؤول إلى الصفر مع زيادة n (بالنسبة لـ MSE $(\hat{\theta})$، ارجع إلى المناقشة التي قدمناها من قبل في هذه النقطة).

مثال 26

دع X_1، X_2، . . . ، X_n تمثل عينة عشوائية من توزيع ما، له توقع μ وتباين σ^2. اثبت أن متوسط العينة $\bar{X}$ هو مقدر متسق لـ μ.

مع الإحصاء الأولي معروف أن $E(\bar{X}) = \mu$ و $\text{Var}(\bar{X}) = \sigma^2/n$ وبما أن $E(\bar{X}) = \mu$ بغض النظر عن حجم العينة فهو مقدر غير متحيز. عمومًا مع زيادة حجم العينة فإن $\text{Var}(\bar{X})$ يتجه ناحية الصفر. وبالتالي $\bar{X}$ هو مقدر متسق لـ μ.

القواعد التالية للنهايات الاحتمالية تحتاج إلى تفصيل كالتالي:

1- الثبات (خاصية Slutsky). إذا كان $\hat{\theta}$ مقدر متسق لـ θ، وإذا كان $h(\hat{\theta})$ دالة متصلة في $\hat{\theta}$ فإن:

$$\operatorname*{plim}_{n\to\infty} h(\hat{\theta}) = h(\theta)$$

(7) بشكل فني أكثر فإن: $\lim_{n\to\infty} \text{var}(\hat{\theta}_n) = \theta$ و $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$

مايعنيه ذلك إنه إذا كان $\hat{\theta}$ مقدر متسق لـ θ، فإن $1\backslash\hat{\theta}$ هو أيضًا مقدر متسق لـ $1\backslash\hat{\theta}$، ويكون أيضًا $\text{Log}(\hat{\theta})$ مقدر متسق لـ $\text{Log}(\hat{\theta})$. لاحظ أن هذه الخاصية لاتتحقق مع التوقع E، بمعنى إذا كانت $\hat{\theta}$ مقدر غير متحيز لـ θ [بمعنى أن: $E(\hat{\theta}) = \theta$] فإن ذلك لايعني أن $1\backslash\hat{\theta}$ مقدر غير متحيز لـ $1\backslash\hat{\theta}$ أي أن: $E(1/\hat{\theta}) \neq 1/E(\hat{\theta}) \neq 1/\theta$

2 - إذا كان b ثابتًا فإن:

$$\underset{n\to\infty}{\text{plim}}\, b = b$$

أي أنها النهاية الاحتمالية لثابت هي نفس هذا الثابت.

3 - إذا كان $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدرات متسقة فإن:

$$\text{plim}\,(\hat{\theta}_1 + \hat{\theta}_2) = \text{plim}\,\hat{\theta}_1 + \text{plim}\,\hat{\theta}_2$$

$$\text{plim}\,(\hat{\theta}_1\hat{\theta}_2) = \text{plim}\,\hat{\theta}_1\ \text{plim}\,\hat{\theta}_2$$

$$\text{plim}\left(\frac{\hat{\theta}_1}{\hat{\theta}_2}\right) = \frac{\text{plim}\,\hat{\theta}_1}{\text{plim}\,\hat{\theta}_2}$$

الخاصيتان الأخيرتان، في العموم، لايتحققان مع علامة التوقع E. بمعنى أن $E(\hat{\theta}_1/\hat{\theta}_2) \neq E(\hat{\theta}_1)/E(\hat{\theta}_2)$ وبالمثل $E(\hat{\theta}_1\,\hat{\theta}_2) \neq E(\hat{\theta}_1)E(\hat{\theta}_2)$. إذا كان $\hat{\theta}_1$ و $\hat{\theta}_2$ مستقلين، فإن $E(\hat{\theta}_1\,\hat{\theta}_2) = E(\hat{\theta}_1)\,E(\hat{\theta}_2)$ كما سبق ولاحظنا ذلك.

الكفاية التقاربية Asymptotic Efficiency: إذا كان $\hat{\theta}$ مقدر لـ θ. تباين التوزيع التقاربي لـ $\hat{\theta}$ يقال عنه تباين تقاربي لـ $\hat{\theta}$. إذا كان $\hat{\theta}$ متسقًا وتباينه التقاربي أصغر من التباين التقاربي. لباقي مقدرات θ المتسقة، فإن $\hat{\theta}$ يقال إن له كفاية تقاربية.

التوزيع الطبيعي التقاربي Asymptotic Normality: يقال إن $\hat{\theta}$ له توزيع طبيعي تقاربي إذا كان توزيع العينة الخاص به يؤول إلى التوزيع الطبيعي كلما يزداد حجم العينة. على سبيل المثال، النظرية الإحصائية تقول بأنه إذا كان X_1، X_2، . . . ، X_n لها توزيع طبيعي وهي متغيرات مستقلة لها توقع μ وتباين σ^2 فإن متوسط العينة $\overline{X}$ يتبع هو أيضًا التوزيع الطبيعي بتوقع μ وتباين $\frac{\sigma^2}{n}$ في العينات صغيرة الحجم وكبيرة الحجم أيضًا. ولكن إذا كان X_i مستقلين وتوقعها μ وتباينها σ^2 ولكن ليس بالضرورة لها التوزيع الطبيعي، فإن توقع العينة $\overline{X}$ له التوزيع الطبيعي تقاربيًا بتوقع μ وتباين σ^2/n أي أنه كلما يزداد حجم العينة n فإن توقع العينة يؤول إلى التوزيع الطبيعي بتوقع μ وتباين σ^2/n. وهذا في الحقيقة هو نظرية النزعة المركزية التي تم استعراضها من قبل.

8.A الاستدلال الإحصائي.. اختبارات الفروض :
STATISTICAL INFERENCE: HYPOTHESIS TESTING

التقدير واختبارات الفروض يعتبران طريقتين توأمتين في الاستدلال الإحصائي التقليدي. بعد أن استعرضنا مشكلة التقدير دعنا الآن باختصار نستعرض مشكلة اختبارات الفروض.

مشكلة اختبارات الفروض يمكن عرضها كالتالي: افترض أن لدينا X rv لـه PDF $f(x;\theta)$ معروفة، حيث θ هو معلمة التوزيع. افترض أننا حصلنا على عينة عشوائية من الحجم n وحصلنا على تقدير بنقطة لـθ وهو $\hat{\theta}$ بما أن θ الحقيقية غير معلومة، فهنا يظهر هذا السؤال: هو المقدر $\hat{\theta}$ "متماشي" مع قيمة افتراضية ما لـθ مثلاً $\theta=\theta^*$ حيث θ^* قيمة رقمية محددة لـθ؟ بمعنى آخر، هل هذه العينة جاءت من PDF، $f(x,\theta)=\theta^*$؟. باستخدام لغة اختبارات الفروض، فإن $\theta=\theta^*$ يطلق عليه الفرض العدمي، ويرمز له عمومًا بالرمز H_0. الفرض العدمي يتم اختباره ضد الفرض البديل، والذي يرمز له عادة H_1، والذي قد يكون مثلاً $\theta\neq\theta^*$. (لاحظ أنه: في بعض الكتب H_0 و H_1 يرمز لهما H_1 و H_2 بالترتيب).

الفرض العدمي والفرض البديل، قد يكونان بسيطين أو مركبين. يقال إن الفرض البسيط إذا حدد قيمة أو مجموعة من القيم للمعلمة أو للمعالم المجهولة، وبخلاف ذلك يسمى فرضًا مركبًا. وبالتالي إذا كانت $X\sim N(\mu,\sigma^2)$ وكان لدينا:

$$H_0: \mu=15 \quad \text{و} \quad \sigma=2$$

فهذا فرض بسيط، أما :

$$H_0: \mu=15 \quad \text{و} \quad \sigma>2$$

يسمى فرضًا مركبًا، حيث إن قيمة σ لم يتم تحديدها بالضبط.

لاختبار الفرض العدمي (أي مدى صلاحيته) فإننا تستخدم معلومات العينة للحصول على ما يسمى إحصاء الاختبار. عادة هذا الإحصاء يعتمد على التقدير بنقطة للمعلمة المجهولة. ثم نحاول الحصول على التوزيع الاحتمالي لهذا الإحصاء، وتستخدم فترة الثقة واختبار المعنوية لاختبار الفرض العدمي. هذه الطريقة يمكن شرحها كالتالي:

لتحديد الأفكار، دعنا نستخدم مثال 23، والمتعلق بأطوال الرجال (X) في المجتمع، قد سبق وذكرنا أن:

$$X_i \sim N(\mu, \sigma^2) = N(\mu, 2.5^2)$$
$$\bar{X} = 67 \qquad n = 100$$

دعنا نفترض التالي :

$$H_0: \mu = \mu^* = 69$$
$$H_1: \mu \neq 69$$

السؤال الآن : هل يمكن للعينة ذات $\bar{X} = 67$ ، إحصاء الاختبار، أن تأتي من مجتمع له قيمة متوسط تساوي 69 ؟ واضح أننا قد لا نستطيع رفض الفرض العدمي إذا كان $\bar{X}$ "قريبة بشكل كاف" لـ μ^* وبخلاف ذلك يمكن رفض الفرض العدمي لصالح الفرض البديل. ولكن كيف نقرر إذا كانت $\bar{X}$ "قريبة بدرجة كافية" من μ^* أم لا؟ يمكن أن تستخدم إحدى الطريقتين التاليتين : (1) فترات الثقة و (2) اختبار المعنوية، كل منهما سيؤدي إلى نفس الاستنتاج عندما يتطبق على نفس المشكلة أو الموضوع.

طريقة فترة الثقة : The Confidence Interval approach

بما أن $X_i \sim N(\mu, \sigma^2)$ فإننا نعرف أن الإحصاء $\bar{X}$ موزع كالتالي :

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

وحيث إننا نعلم التوزيع الاحتمالي لـ $\bar{X}$ ، لماذا لانكون مثلاً $100(1-\alpha)$ فترة ثقة لـ μ معتمدة على $\bar{X}$ ، ونرى ما إذا كانت هذه الفترة تحتوي على $\mu = \mu^*$ أم لا؟ إذا كانت تحتوي على القيمة μ^*، فإننا قد لا نستطيع رفض الفرض العدمي. أما إذا كانت لاتحتوي على القيمة μ^* فإننا نرفض الفرض العدمي. وبالتالي، إذا كانت $\alpha = 0.05$ سيكون لدينا 95% فترة ثقة، وإذا كانت هذه الفترة تحتوي على μ^* فإننا لانرفض الفرض العدمي، حيث 95 فترة ثقة من الـ 100 فترة التي يتم تكوينها ستحتوي على μ^* .

الطريقة بالفعل تتم كالتالي :

بما أن : $\bar{X} \sim N(\mu, \sigma^2/n)$ فإن :

$$Z_i = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

وهذا هو التوزيع الطبيعي القياسي. وبالتالي من جدول التوزيع الطبيعي القياسي نعرف أن :

$$\Pr(-1.96 \leq Z_i \leq 1.96) = 0.95$$

أي أن :

$$\Pr\left(-1.96 \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq 1.96\right) = 0.95$$

والذي يساوي عند إعادة ترتيب مقاديره التالي :

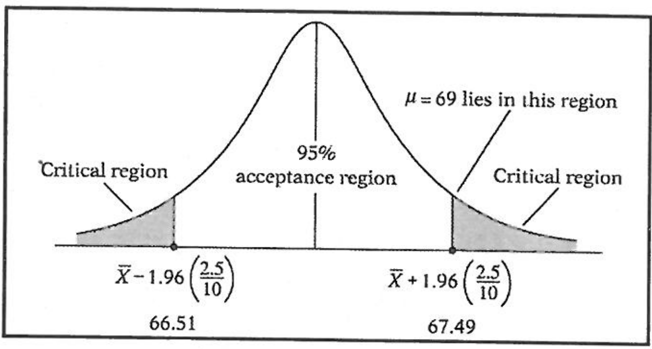
$$\Pr\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right] = 0.95$$

وهذا هو 95% فترة ثقة لـ μ. بمجرد تكوين تلك الفترة يصبح اختبار الفرض العدمي أمرًا في منتهى البساطة. كل ما نحتاج إلى عمله هو تحديد ما إذا كانت هذه الفترة تحتوي على $\mu = \mu^*$ أم لا. إذا كانت تحتوي على تلك القيمة، فإننا لانستطيع رفض الفرض العدمي، أما إذا كانت لاتحتوي على تلك القيمة، فإننا نرفض الفرض العدمي. بالعودة إلى المثال، فقد كونا بالفعل الـ 95% فترة ثقة لـ μ وهي :

$$66.51 \leq \mu \leq 67.49$$

من الواضح أن هذه الفترة لاتحتوي على النقطة $\mu = 69$، وبالتالي تستطيع أن ترفض الفرض العدمي القائل بأن μ الحقيقية تساوي 69 بمعامل ثقة 95% . هندسيًا هذا الوضع مشروح في الشكل (12.A).

باستخدام مصطلحات اختبارات الفروض، فترة الثقة التي تم تكوينها، تسمى منطقة القبول، والمناطق التي تقع خارجها تسمى المنطقة الحرجة أو منطقة الرفض للفرض العدمي. الحدود الدنيا والعليا لمنطقة القبول (والتي تفصلها عن منطقة الرفض) تسمى القيم الحرجة. وبالتالي بلغة اختبارات الفروض، إذا وقعت القيمة الافتراضية داخل منطقة القبول، لاتستطيع رفض الفرض العدمي، وبخلاف ذلك نرفض الفرض العدمي. من المهم ملاحظة أنه عند اتخاذ قرار رفض أو عدم رفض H_0، فإننا من المحتمل أن نواجه أحد نوعين من الأخطاء التالية : (1) نرفض H_0 في حين أنه في الحقيقة سليم، ويسمى ذلك الخطأ من النوع I (فمثلاً في المثال السابق $\bar{X} = 67$ قد تكون مسحوبة من مجتمع قيمة المتوسط الفعلية 69) أو (2) قد لانرفض الفرض العدمي، في حين أنه في الحقيقة خطأ ويسمى ذلك الخطأ من النوع II).



شكل (12.A) 95% فترة ثقة لـ μ

وبالتالي اختبار الفرض العدمي لايحقق القيمة الصحيحة لـ μ. ولكنه يعطي معنى لتقرير ماتستطيع فعله كما لو أن $\mu = \mu^*$.

الأخطاء من النوع I والنوع II. بشكل تنظيمي لدينا التالي:

القـــرار	الواقع الحقيقي	
	H_0 سليم	H_0 خطأ
رفض H_0	الخطأ من النوع I	لايوجد خطأ
عدم رفض H_0	لايوجد خطأ	الخطأ من النوع II

في أفضل الأحوال، نحن نريد تقليل كل من أخطاء النوع I والنوع II. ولكن للأسف بالنسبة لحجم عينة محدد، لاتستطيع تقليل نوعين من الخطأ معًا آنيًا. الطريقة التقليدية لهذه المشكلة، والتي ظهرت في أبحاث Neyman و Pearson، هي أن نفترض أن الخطأ من النوع I أكثر خطورة عمليًا من الخطأ من النوع II. وبالتالي لابد أن يحاول الفرد جعل احتمال حدوث الخطأ من النوع I أقل ما يمكن مثل 0.01 أو 0.05 ثم يحاول بعد ذلك تقليل احتمال حدوث الخطأ من النوع II بقدر المستطاع.

عرف سابقًا، أن احتمال حدوث الخطأ من النوع I هو α، وتسمى مستوى المعنوية. واحتمال حدوث الخطأ من النوع II β. واحتمال عدم حدوث الخطأ من النوع II يسمى قوة الاختبار. بعبارة أخرى، قوة الاختبار هي القدرة على رفض فرض عدمي خاطئ. الطريقة التقليدية لاختبارات الفروض تعتمد على تحديد مستوى معين مثل 0.01 (أو 1%) أو 0.05 (أو 5%) لـ α ثم نحاول تعظيم قوة الاختبار، أي تقليل β.

من الضروري أن يستوعب القارئ مفهوم قوة الاختبار، وذلك المفهوم نستعرضه حاليًا بالمثال التالي[8].

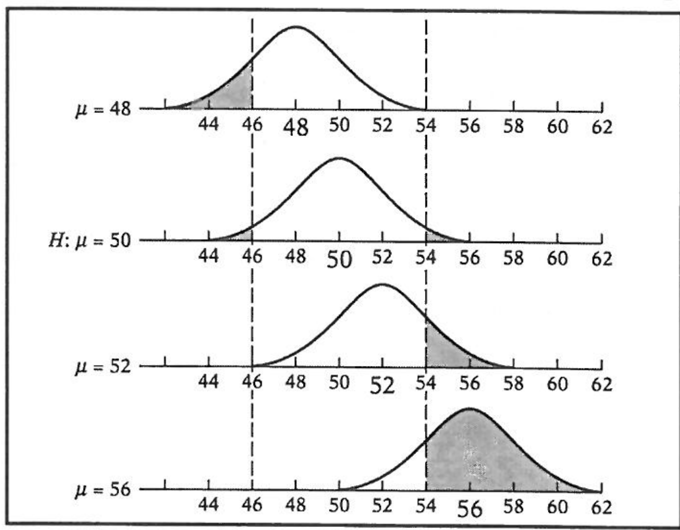
افترض أن $X \sim N(\mu, 100)$، أي أن X له التوزيع الطبيعي بتوقع μ وتباين 100. افترض أن $\alpha = 0.05$ وأن حجم العينة 25 مفردة والتي تستخدمها لحساب متوسط العينة $\bar{X}$. افترض أننا نريد اختبار الفرض: $H_0: \mu = 50$ بما أن X له التوزيع الطبيعي، فإننا نعرف أن متوسط العينة يتبع أيضًا التوزيع الطبيعي كالتالي: $\bar{X} \sim N(\mu, 100/25)$ وبالتالي بافتراض صحة الفرض العدمي القائل بأن $\mu = 50$، فإن الـ 95% فترة ثقة لـ μ هو $(\bar{X} \pm 1.96\sqrt{100/25}) = \bar{X} \pm 3.92$، وذلك يساوي (53.92 و 46.08)، وبالتالي المنطقة الحرجة تتكون من كل قيم $\bar{X}$ الأقل من 46.08 أو الأكبر من 53.92 وبالتالي سنرفض الفرض العدمي بأن المتوسط الحقيقي 50 إذا وجد متوسط العينة أقل من 46.08 أو أكبر من 53.92.

ولكن ما احتمال أن $\bar{X}$ تقع في المنطقة الحرجة السابقة إذا كانت μ الحقيقية لها قيمة مختلفة عن 50؟ افترض أن لدينا الفروض البديلة التالية: $\mu = 48$، $\mu = 52$ و $\mu = 56$. إذا كان أحد هذه البدائل صحيحًا، ستكون هذه القيمة هي قيمة المتوسط الحقيقية لتوزيع $\bar{X}$. الخطأ القياسي لن يتغير بالنسبة للبدائل الثلاثة السابقة، حيث إن σ^2 مازال مفترضًا أنه يساوي 100.

المنطقة المظللة في شكل (13.A) توضح احتمال أن تقع $\bar{X}$ في المنطقة الحرجة إذا كان الفرض البديل صحيحًا. ويمكنك التأكد من أن هذا الاحتمال يساوي 0.17 (لـ $\mu = 48$) و0.05 (لـ $\mu = 50$) و 0.17 (لـ $\mu = 52$) و 0.85 (لـ $\mu = 56$). وكما ترى في الشكل، عندما يكون هناك فرق حقيقي بين القيمة الحقيقية لـ μ، والقيمة المفترضة (والتى تساوي هنا $\mu = 50$) فإن احتمال رفض الفرض العدمي يزداد، ولكن عندما لاتختلف كثيرًا القيمة الحقيقية عن القيمة المفترضة الموجودة في الفرض العدمي، فإن احتمال الرفض يكون صغيرًا. وبالطبع لابد أن يكون ذلك منطقيًا إذا كان الفرض العدمي والبديل قريبًا لبعضهما البعض. ويمكن فهم ذلك بصورة أفضل إذا نظرت إلى الشكل (14.A) ويمثل رسمًا لدالة القوة والمنحنى الموجود بالشكل يسمى منحنى القوة.

(8) المناقشة التالية والأشكال البيانية معتمدة على
Helen M. Walker and Joseph Lev, Statistical Inference, Holt, Rinehart and Winston, New York, 1953, pp. 161-162.

وسيدرك القارئ الآن، أن معامل الفترة (α– 1) الذي ناقشناه سابقًا هو ببساطة واحد مطروح منه احتمال وقوع الخطأ من النوع I.

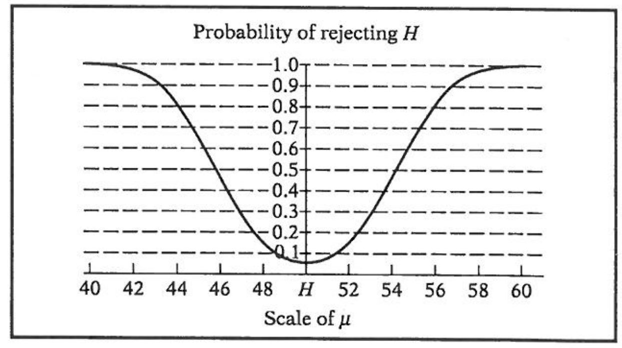


شكل (13.A) توزيع X عندما $N= 25$ ، $\sigma= 10$ و 56 أو $\mu= 48, 50, 25$ في حالة $H_0: \mu= 50$ ، المنطقة الحرجة لـ $\alpha= 0.05$ هي $\bar{X} < 46.1$ و $53.9 < \bar{X}$. المنطقة المظللة تعبر عن احتمال أن $\bar{X}$ تقع داخل المنطقة الحرجة هذا الاحتمال هو :

0.17 إذا كان $\mu= 48$ 0.17 إذا كان $\mu= 52$

0.05 إذا كان $\mu= 50$ 0.85 إذا كان $\mu= 56$

وبالتالي 95% فترة ثقة، تعني أننا على استعداد لقبول على الأكثر 5% احتمال لوقوع الخطأ من النوع I، فنحن لانريد رفض فرض صحيح بأكثر من خمس مرات من كل 100 مرة.



شكل (14.A) دالة القوة لاختبار الفرض $\mu = 50$ عند $N = 10$ ، $\sigma = 10$ و $\alpha = 0.05$

قيمة P أو مستوى المعنوية التام : The P - Value or Exact level of significance

بدلاً من الاختبار المسبق لـ α عند أي مستوى معين مثل 1، 5 أو10% ، فإنه يمكن الحصول على قيمة P (الاحتمال) أو مستوى المعنوية التام لإحصاء ما قيمة P تعرف على أنها أقل مستوى معنوية يمكن عند رفض الفرض العدمي.

افترض أنه في مسألة ما فيها 20 درجة حرية، وقيمة t تساوي 3.552 قيمة P أو الاحتمال التام للحصول على قيمة t تساوي 3.552 أو أكثر يمكن حسابه من جدول (2.D) وسيساوي 0.001 (اختيار طرف واحد) أو 0.002 (اختيار ذي طرفين). يمكننا القول بأن قيمة t المحسوبة والمساوية لـ 3.552 لها معنوية إحصائية عند 0.001 أو 0.002 معتمدين على ما إذا كان الاختيار ذا طرف واحد أو ذا طرفين.

هناك العديد من حزم البرامج الإحصائية التي تقوم بحساب قيمة P للإحصاء المقدر. وبالتالي ينصح القارئ بأن يستخدم قيمة P عندما يكون ذلك متاحاً.

طريقة اختبار المعنوية : The Test of Significance Approach

تذكر أن :

$$Z_i = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

في أي تطبيق، $\bar{X}$ و n يكونان معروفين (أو ممكن تقديرهما)، ولكن قيمة μ الحقيقية و σ غير معروفين. ولكن إذا كانت σ محددة يمكن أن نفترض (تحت صحة الفرض العدمي) بأن $\mu = \mu^*$، قيمة رقمية محددة، وبالتالي Z_i يمكن أن يتم حسابه مباشرة، ويمكن بسهولة البحث في جدول التوزيع الطبيعي لحساب احتمال الحصول على قيمة Z المحسوبة. إذا كان هذا الاحتمال صغيراً مثلاً أقل من 5% أو 1%، فإننا نرفض الفرض العدمي ،إذا كان الفرض العدمي سليماً فإن فرصة الحصول على قيمة Z المحددة لابد أن يكون كبيراً. هذه هي الفكرة العامة وراء طريقة اختيار المعنوية لاختبارات الفروض. الفكرة الأساسية هي إحصاء الاختيار (هنا هي إحصاء Z) وتوزيعه الاحتمالي بافتراض صحة $\mu = \mu^*$.

يتضح أنه في الحالة الحالية، فإن الاختيار معروف باسم اختيار Z، حيث إننا نستخدم قيمة Z (التوزيع الطبيعي القياسي).

بالعودة إلى مثالنا الحالي، إذا كان $\mu = \mu^* = 69$ فإن إحصاء Z يصبح

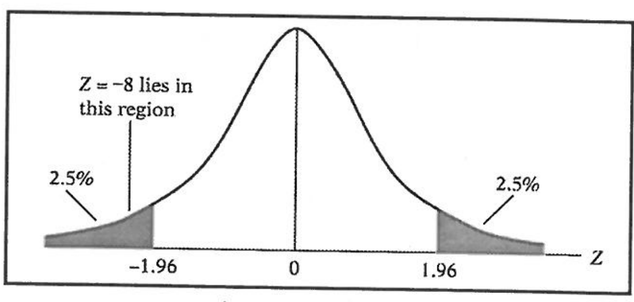
$$Z = \frac{\bar{X} - \mu^*}{\sigma/\sqrt{n}}$$

$$= \frac{67 - 69}{2.5/\sqrt{100}}$$

$$= -2/0.25 = -8$$

إذا نظرنا في جدول التوزيع الطبيعي (1.D) سنجد أن احتمال الحصول على قيمة Z مثل السابقة هو احتمال صغير جداً (لاحظ أن: احتمال أن قيمة Z تزيد على 3 أو 3- هو حوالي 0.001. وبالتالي احتمال أن تزيد Z على 8 سيكون أصغر) وبالتالي فإنه يمكن رفض الفرض العدمي القائل أن $\mu = 69$ بمعلومية ذلك احتمال الحصول على $\bar{X}$ تساوي 67 صغير جداً مما يجعل هناك شك في أن هذه العينة جاءت من مجتمع قيمته المتوسطة هي 69. بيانيًا هذه الحالة موضحة في شكل (15.A).

وباستخدام لغة اختيار المعنوية، فإننا نقول إن الاختيار (الإحصاء) معنوي إذا كنا نقصد بوجه عام إمكانية رفض الفرض العدمي. ويعتبر الإحصاء معنويًا إذا كان احتمال حصولنا عليه أقل من أو يساوي α، احتمال وقوع الخطأ من النوع I. وبالتالي إذا كانت $\alpha = 0.05$ نعرف أن احتمال الحصول على قيمة لـ Z تساوي 1.96- أو 1.96 هو 5% (أو 2.5% في كل طرف للتوزيع الطبيعي القياسي). في مثالنا التوضيحي قيمة Z كانت 8-. وبالتالي احتمال الحصول على قيمة Z أقل بكثير من 2.5% مما تجعلها أقل من الاحتمال المحدد سابقًا لحدوث الخطأ من النوع I، ولهذا السبب فإن قيمة Z المحسوبة $Z = -8$ معنوية إحصائيًا، أي أننا نرفض الفرض العدمي بأن μ^* هي 69. بالطبع سنصل إلى نفس الاستنتاج باستخدام طريقة فترة الثقة لاختبارات الفروض.



شكل (15.A) توزيع إحصاء Z

والآن دعنا نلخص الخطوات المطلوبة لاختيار أي فرض إحصائي :

الخطوة 1 - حدد الفرض العدمي H_0 والفرض البديل H_1

(مثلاً: $H_0: \mu = 69$ و $H_1: \mu \neq 69$)

الخطوة 2 - اختر إحصاء الاختيار (مثلاً $\bar{X}$)

الخطوة 3 - حدد التوزيع الاحتمالي للإحصاء (مثلاً $\bar{X} \sim N(\mu, \sigma^2/n)$)

الخطوة 4 - اختر مستوى المعنوية (أي احتمال حدوث الخطأ من النوع I) α .

الخطوة 5 - باستخدام التوزيع الاحتمالي للإحصاء، كون فترة ثقة $100(1-\alpha)\%$ إذا كانت قيمة المعلمة تحت صحة الفرض العدمي (مثلاً $\mu = \mu^* = 69$) تقع داخل هذه الفترة، منطقة القبول، فإننا لانرفض الفرض العدمي. ولكن إذا وقعت خارجها (بمعنى أنها وقعت في منطقة الرفض) فإننا نرفض الفرض العدمي. ضع في الاعتبار أنك سواء في حالة رفض أو عدم رفض الفرض العدمي، هناك احتمال أن تكون على خطأ $\alpha\%$ في كل مرة.

REFERENCES المراجع :

لمزيد من التفاصيل في المواضيع التي تم مناقشتها في هذا الملحق، يمكن للقارئ أن يستخدم أيًا من المراجع التالية :

Hoel, Paul G.: Introduction to Mathematical Statistics, 4th ed., John Wiley & Sons, New York, 1974. This book provides a fairly simple introduction to various aspects of mathematical statistics.

Freund, John E., and Ronald E. Walpole: Mathematical Statistics, 3d ed., Prentice Hall, Englewood Cliffs, N.J., 1980. Another introductory textbook in mathematical statistics.

Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes: Introduction to the Theo?y of Statistics, 3d ed., McGraw-Hill, New York, 1974. This is a comprehensive introduction to the theory of statistics but is somewhat more difficult than the preceding two textbooks.

Newbold, Paul: Statistics for Business and Economics, Prentice Hall, Englewood Cliffs, N.J., 1984. A comprehensive nonmathematical introduction to statistics with lots of worked-out problems.

ملحق B

مبادئ جبر المصفوفات
RUDIMENTS OF MATRIX ALGEBRA

هذا الملحق يستعرض أساسيات جبر المصفوفات اللازمة لفهم ملحق C، وبعض الموضوعات الموجودة في الفصل (18). المناقشة ستكون غير عميقة، ولاتوجد أي إثباتات نظرية. للمزيد من التفاصيل وللإثباتات النظرية، يمكن للقارئ أن يلجأ إلى المراجع.

1.B تعريفات : DEFINITIONS

المصفوفة : Matrix

المصفوفة هي منظومة مستطيلة من مجموعة من الأرقام أو العناصر الموجودة في صفوف وأعمدة. لنكن أكثر تحديداً، فالمصفوفة من الدرجة أو البعد M و N (تكتب $M \times N$) هي عبارة عن $M \times N$ عنصران موجودان في M صف و N عمود. وبالتالي إذا اعتبرنا الحروف سوداء الطباعة ترمز للمصفوفات، فإن مصفوفة $\mathbf{A}\ (M \times N)$، يمكن أن تكتب كالتالي:

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{M1} & a_{M2} & a_{M3} & \cdots & a_{MN} \end{bmatrix}$$

حيث a_{ij} هو العنصر الذي يظهر في الصف رقم i، والعمود رقم j من $\mathbf{A}$ و $[a_{ij}]$ ترمز إلى المصفوفة $\mathbf{A}$ التي عناصرها هي a_{ij}. درجة أو بعد المصفوفة عبارة عن عدد الصفوف، وعدد الأعمدة، وعادة يكتب تحت المصفوفة لتسهيل معرفته.

$$\underset{2\times 3}{\mathbf{A}} = \begin{bmatrix} 2 & 3 & 5 \\ 6 & 1 & 3 \end{bmatrix} \qquad \underset{3\times 3}{\mathbf{B}} = \begin{bmatrix} 1 & 5 & 7 \\ -1 & 0 & 4 \\ 8 & 9 & 11 \end{bmatrix}$$

الثابت **Scalar**: هو عبارة عن رقم حقيقي مفرد، وبالتالي هو مصفوفة 1×1.

المتجه العمودي Column Vector : المصفوفة المكونة من M صف وعمود واحد فقط، تسمى متجه عمودي. وباعتبار الحروف سوداء الطباعة تعبر عن المتجهات، فإن التالي هو مثال للمتجه العمودي.

$$\underset{4\times1}{\mathbf{x}} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 9 \end{bmatrix}$$

المتجه الصفي : Row Vector

المصفوفة المكونة من صف واحد فقط، و N عمود تسمى متجهًا صفيًا.

$$\underset{1\times4}{\mathbf{x}} = [1 \quad 2 \quad 5 \quad -4] \qquad \underset{1\times5}{\mathbf{y}} = [0 \quad 5 \quad -9 \quad 6 \quad 10]$$

التدوير : Transposition

مدور المصفوفة $\mathbf{A}$ ذات الأبعاد $M \times N$ يرمز له بالرمز $\mathbf{A}'$ (وتقرأ $\mathbf{A}$ المشروطة أو مدور $\mathbf{A}$) وهو عبارة عن مصفوفة $N \times M$، حيث يتم تبديل الصفوف والأعمدة الخاصة بـ $\mathbf{A}$، بمعنى أن الصف i في $\mathbf{A}$ يصبح هو العمود i في $\mathbf{A}'$. فمثلاً:

$$\underset{3\times2}{\mathbf{A}} = \begin{bmatrix} 4 & 5 \\ 3 & 1 \\ 5 & 0 \end{bmatrix} \qquad \underset{2\times3}{\mathbf{A}'} = \begin{bmatrix} 4 & 3 & 5 \\ 5 & 1 & 0 \end{bmatrix}$$

وبما أن المتجه هو حالة خاصة من المصفوفة، فإن مدور متجه صفي هو متجه عمودي، ومدور متجه عمودي هو متجه صفي. وبالتالي:

$$\mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad و \qquad \mathbf{x}' = [4 \quad 5 \quad 6]$$

وسنرمز دائمًا للمتجه الصفي بالمتجه المدور (المشروط).

المصفوفة الجزيئية : Submatrix

بالنسبة لأي مصفوفة $\mathbf{A}$ $M \times N$، إذا تم إلغاء كل الأعمدة باستثناء s عمود وكل الصفوف باستثناء r صف، فإن المصفوفة الناتجة من الدرجة $r \times s$ تسمى مصفوفة جزيئية من $\mathbf{A}$. وبالتالي إذا كان:

$$\underset{3\times3}{\mathbf{A}} = \begin{bmatrix} 3 & 5 & 7 \\ 8 & 2 & 1 \\ 3 & 2 & 1 \end{bmatrix}$$

إذا حذفنا الصف الثالث والعمود الثالث من A، نحصل على:

$$\underset{2\times2}{\mathbf{B}} = \begin{bmatrix} 3 & 5 \\ 8 & 2 \end{bmatrix}$$

وهي مصفوفة جزيئية من A من الدرجة 2×2.

2.B أنواع المصفوفات : TYPES OF MATRICES

المصفوفة المربعة : Square Matrix

المصفوفة التي يكون لها نفس العدد من الصفوف والأعمدة تسمى مصفوفة مربعة.

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & 5 & 8 \\ 7 & 3 & 1 \\ 4 & 5 & 0 \end{bmatrix}$$

المصفوفة القطرية : Diagonal Matrix

المصفوفة المربعة التي لها على الأقل رقم واحد على القطر الرئيسي لايساوي الصفر (القطر الرئيسي يبدأ من أعلى شمال المصفوفة حتى الركن الأسفل على ناحية اليمين) وباقي العناصر تساوي الصفر يطلق عليها مصفوفة قطرية.

$$\underset{2\times2}{\mathbf{A}} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \qquad \underset{3\times3}{\mathbf{B}} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

المصفوفة الثابتة : Scalar Matrix

المصفوفة القطرية التي تكون عناصر القطر الرئيسي فيها كلها متساوية، تسمى المصفوفة الثابتة. وكمثال على المصفوفة القطرية مصفوفة التباين- التغاير للمجتمع مقادير الأخطاء العشوائية في نموذج الانحدار الخطي التقليدي المعطى في المعادلة (3.2.C) وهي كالتالي:

$$\text{var-cov}(\mathbf{u}) = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 \end{bmatrix}$$

مصفوفة الوحدة : Identity or Unit Matrix

المصفوفة القطرية التي تكون عناصر قطرها الرئيسي كلها تساوي الـ 1، تسمى مصفوفة الوحدة ويرمز لها بالرمز **I**. وهي تعتبر حالة خاصة من المصفوفة الثابتة.

$$\underset{3\times3}{\mathbf{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \underset{4\times4}{\mathbf{I}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

المصفوفة المتماثلة : Symmetric Matrix

المصفوفة المربعة التي تكون العناصر الموجودة فيها أعلى القطر الرئيسي هي نسخة متطابقة للعناصر الموجودة أسفل القطر الرئيسي، تسمى مصفوفة متماثلة. بعبارة أخرى، المصفوفة المتماثلة يكون مدورها يساوي نفس المصفوفة الأصلية، أي أن $\mathbf{A} = \mathbf{A}'$ وأن العنصر a_{ij} في **A** يساوي العنصر a_{ji} في $\mathbf{A}'$. كمثال لذلك مصفوفة التباين- التغاير المعطاة في المعادلة (2.2C) ومثال آخر هو مصفوفة الارتباط المعطاة في (1.5.C).

المصفوفة الصفرية : Null Matrix

المصفوفة التي تكون جميع عناصرها تساوي الصفر، تسمى مصفوفة صفرية، ويرمز لها بالرمز **0**.

المتجه الصفري : Null Vector

المتجه سواء عمودي أو صفي، الذي تكون جميع عناصره تساوي الصفر، يسمى متجه صفري، ويرمز له أيضًا بالرمز **0**.

المصفوفات المتساوية : Equal Matrices

المصفوفتان **A** و **B** يقال عنهما إنهما متساويتان إذا كانتا من نفس الدرجة، وعناصرهما المتناظرة متساوية، أي أن $a_{ij} = b_{ij}$ لكل i و j، مثلاً المصفوفتان:

$$\underset{3\times3}{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 5 \\ 0 & -1 & 2 \\ 5 & 1 & 3 \end{bmatrix} \quad \text{و} \quad \underset{3\times3}{\mathbf{B}} = \begin{bmatrix} 3 & 4 & 5 \\ 0 & -1 & 2 \\ 5 & 1 & 3 \end{bmatrix}$$

متساويتان، أي أن $\mathbf{A} = \mathbf{B}$.

3.B عمليات على المصفوفات : MATRIX OPERATIONS

جمع المصفوفات : Matrix Addition

إذا كانت $\mathbf{A} = [a_{ij}]$ و $\mathbf{B} = [b_{ij}]$. إذا كانت A و B لهما نفس الدرجة، فإننا نعرف جمع المصفوفتين كالتالي:

$$\mathbf{A} + \mathbf{B} = \mathbf{C}$$

حيث C هي مصفوفة من نفس درجة A و B، ويتم الحصول عليها من $c_{ij} = a_{ij} + b_{ij}$ لكل i و j، أي أن C نحصل عليها بجمع العناصر المتناظرة لـ A مع B. إذا كان هذا الجمع ممكن القيام به فتسمى A و B مصفوفات متناسبتان للجمع. مثلاً إذا كان:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \end{bmatrix} \quad \text{و} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & -1 & 3 \\ -2 & 0 & 1 & 5 \end{bmatrix}$$

و $\mathbf{C} = \mathbf{A} + \mathbf{B}$ هي

$$\mathbf{C} = \begin{bmatrix} 3 & 3 & 3 & 8 \\ 4 & 7 & 9 & 14 \end{bmatrix}$$

طرح المصفوفات : Matrix Subtraction

طرح المصفوفات مماثل لجمع المصفوفات، باستثناء أن $\mathbf{C} = \mathbf{A} - \mathbf{B}$، أي أننا نطرح عناصر B من العناصر المناظرة في A للحصول على C، علمًا بأن A و B لهما نفس الدرجة.

الضرب في ثابت : Scalar Multiplication

لضرب المصفوفة A في الثابت λ (أي رقم حقيقي)، فإننا نضرب كل عنصر في المصفوفة في λ:

$$\lambda\mathbf{A} = [\lambda a_{ij}]$$

فمثلاً إذا كان λ= 2 وإذا كانت:

$$\mathbf{A} = \begin{bmatrix} -3 & 5 \\ 8 & 7 \end{bmatrix}$$

فإن :

$$\lambda\mathbf{A} = \begin{bmatrix} -6 & 10 \\ 16 & 14 \end{bmatrix}$$

ضرب المصفوفات : Matrix Multiplication

إذا كانت A مصفوفة $M \times N$، و B مصفوفة $N \times P$. فإن حاصل ضرب AB (بهذه الدرجات) يعرف على أنه مصفوفة جديدة C من الدرجة $M \times P$ كالتالي:

$$c_{ij} = \sum_{k=1}^{N} a_{ik} b_{kj} \qquad \begin{matrix} i = 1, 2, \ldots, M \\ j = 1, 2, \ldots, P \end{matrix}$$

أي أن عناصر الصف i والعمود j في **C**، يتم الحصول عليهما بضرب عناصر الصف i في **A** مع العناصر المناظرة في العمود j في **B** وجمع كل المقادير معًا. وهذا معروف باسم قاعدة ضرب الصف في العمود لضرب المصفوفات.

وبالتالي للحصول على c_{11}، عناصر الصف الأول والعمود الأول في **C** يتم الحصول عليها بضرب عناصر الصف الأول في **A** مع العناصر المناظرة في العمود الأول في **B**، ثم نجمع كل المقادير. وبالمثل للحصول على c_{12} فإننا نضرب عناصر الصف الأول في **A**، مع العناصر المناظرة في العمود الثاني في **B**، ثم نجمع كل المقادير وهكذا.

لاحظ أنه حتى تكون هناك إمكانية للقيام بعملية ضرب المصفوفات فإن **A** و **B** لابد أن يكونا مناسبين للضرب، بمعنى أن عدد الأعمدة في **A** لابد أن يتساوى مع عدد الصفوف في **B**. فمثلاً:

$$\underset{2\times3}{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 7 \\ 5 & 6 & 1 \end{bmatrix} \quad \text{و} \quad \underset{3\times2}{\mathbf{B}} = \begin{bmatrix} 2 & 1 \\ 3 & 5 \\ 6 & 2 \end{bmatrix}$$

$$\mathbf{AB} = \underset{2\times2}{\mathbf{C}} = \begin{bmatrix} (3\times2)+(4\times3)+(7\times6) & (3\times1)+(4\times5)+(7\times2) \\ (5\times2)+(6\times3)+(1\times6) & (5\times1)+(6\times5)+(1\times2) \end{bmatrix}$$

$$= \begin{bmatrix} 60 & 37 \\ 34 & 37 \end{bmatrix}$$

ولكن إذا كان :

$$\underset{2\times3}{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 7 \\ 5 & 6 & 1 \end{bmatrix} \quad \text{و} \quad \underset{2\times2}{\mathbf{B}} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}$$

فإنه لايمكن تعريف حاصل الضرب **AB**، حيث إن **B** و **A** غير متناسبين للضرب.

خصائص ضرب المصفوفات : Properties of Matrix

1 – ضرب المصفوفات ليس بالضرورة تبادلي، بمعنى أنه عمومًا **AB ≠ BA** وبالتالي ترتيب ضرب المصفوفات يعتبر أمرًا ضروريًا. **AB** يعني أن **A** مضروب أولاً في **B** أو أن **B** مضروبة لاحقًا في **A**.

2 – حتى إذا كان **AB** و **BA** موجودان، فالمصفوفات الناتجة ليست بالضرورة لها نفس الدرجة. فإذا كانت **A** عبارة عن $M \times N$ مصفوفة و **B** عبارة عن $N \times M$ مصفوفة،

فإن **AB** عبارة عن مصفوفة من الدرجة $M \times M$، في حين أن **BA** مصفوفة من الدرجة $N \times N$، وبالتالي لها درجة مختلفة عن المصفوفة الأولى.

3 - حتى إذا كان **A** و **B** كلاهما مصفوفات مربعة وبالتالي **AB** و **BA** كلاهما معرفان، ولكن المصفوفات الناتجة ليست بالضرورة متساوية. فمثلاً:

$$\mathbf{A} = \begin{bmatrix} 4 & 7 \\ 3 & 2 \end{bmatrix} \quad \text{و} \quad \mathbf{B} = \begin{bmatrix} 1 & 5 \\ 6 & 8 \end{bmatrix}$$

فإن:

$$\mathbf{AB} = \begin{bmatrix} 46 & 76 \\ 15 & 31 \end{bmatrix} \quad \text{و} \quad \mathbf{BA} = \begin{bmatrix} 19 & 17 \\ 48 & 58 \end{bmatrix}$$

وبالتالي $\mathbf{BA} \neq \mathbf{AB}$. وكمثال لـ $\mathbf{BA} = \mathbf{AB}$ يمكن أن تكون كل من **A** و **B** مصفوفتا الوحدة.

4 - إذا كان هناك متجه صفي مضروب أولاً في متجه عمودي، فإن حاصل الضرب سيكون ثابتًا. فمثلاً بواقي المربعات الصغرى العادية $\hat{u}_1$، $\hat{u}_2$، . . . ، $\hat{u}_n$، إذا اعتبرنا u هو متجه عمودي و u' هو المتجه الصفي فإن:

$$\begin{aligned}\hat{\mathbf{u}}'\hat{\mathbf{u}} &= [\hat{u}_1 \quad \hat{u}_2 \quad \hat{u}_3 \quad \cdots \quad \hat{u}_n] \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \vdots \\ \hat{u}_n \end{bmatrix} \\ &= \hat{u}_1^2 + \hat{u}_2^2 + \hat{u}_3^2 + \cdots + \hat{u}_n'^2 \\ &= \sum \hat{u}_i^2\end{aligned}$$

انظر المعادلة (8.5C) ثابت

5 - أما إذا ضرب متجه عمودي أولاً في متجه صفي، فإن حاصل الضرب سيكون مصفوفة. وكمثال على ذلك، دعنا نستخدم مقادير أخطاء المجتمع في نموذج الانحدار الخطي التقليدي، $\hat{u}_1$، $\hat{u}_2$، . . . ، $\hat{u}_n$ باعتبار $\boldsymbol{u}$ متجه عمودي و $\boldsymbol{u}'$ متجه صفي فإن لدينا:

$$\begin{aligned}\mathbf{uu}' &= \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{bmatrix} [u_1 \quad u_2 \quad u_3 \quad \cdots \quad u_n] \\ &= \begin{bmatrix} u_1^2 & u_1u_2 & u_1u_3 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & u_2u_3 & \cdots & u_2u_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_nu_1 & u_nu_2 & u_nu_3 & \cdots & u_n^2 \end{bmatrix}\end{aligned}$$

وهذه مصفوفة من الدرجة $n \times n$، لاحظ أن المصفوفة السابقة هي مصفوفة متماثلة.

6 – المصفوفة المضروبة أولاً في متجه عمودي ينتج عنها متجه عمودي.

7 – المتجه الصفي المضروب أولاً في مصفوفة ينتج عنها متجه صفي.

8 – ضرب المصفوفات تجميعي، أي أن $(AB)C = A(BC)$ حيث A مصفوفة $M \times N$ و B مصفوفة $N \times P$ و C مصفوفة $P \times K$.

9 – ضرب المصفوفات توزيعي بالنسبة للجمع، أي أن:

$$A(B + C) = AB + AC \text{ و } (B + C)A = BA + CA$$

تدوير المصفوفة: Matrix Transposition

قد سبق وعرفنا عملية تدوير المصفوفة بتبديل صفوف أو أعمدة المصفوفة معًا (أو المتجه). والآن دعنا نستعرض بعض خصائص التدوير.

1 – تدوير المصفوفة المدورة هو نفسه المصفوفة الأصلية، وبالتالي $(A')' = A$

2 – إذا كان A و B قابلين للجمع، فإن $C = A + B$ و $C' = (A + B)' = A' + B'$، أي أن مدور جمع مصفوفتين يساوي جمع مدورهما.

3 – إذا كانت AB معرفة، فإن $(AB)' = B'A'$، أي أن مدور حاصل ضرب مصفوفتين يساوي حاصل ضرب مدورهما بترتيب عكسي. ولتعميم ذلك فإن: $(ABCD)' = D'C'B'A'$

4 – مدور مصفوفة الوحدة I هو نفسه مصفوفة الوحدة، أي أن: $I' = I$

5 – مدور الثابت هو نفس الثابت، وبالتالي إذا كانت λ ثابتًا فإن $\lambda = \lambda$.

6 – مدور $(\lambda A)'$ هو $\lambda A'$، حيث λ ثابت [لاحظ أن: $(\lambda A)' = A'\lambda' = A'\lambda = \lambda A'$].

7 – إذا كانت A مصفوفة مربعة بحيث إن $A = A'$، فإن A مصفوفة متماثلة (ارجع إلى تعريف المصفوفة المتماثلة والذي تم استعراضه سابقًا).

عكس المصفوفة: Matrix Inversion

معكوس المصفوفة المربعة، والذي يرمز له A^{-1} (ويقرأ معكوس A)، إذا كان موجودًا فإنه عبارة عن مصفوفة مربعة وحيدة، بحيث إن:

$$AA^{-1} = A^{-1}A = I$$

حيث I هي مصفوفة الوحدة التي لها نفس درجة المصفوفة A .

فمثلاً :

$$A = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} -1 & \frac{1}{2} \\ \frac{6}{8} & -\frac{1}{4} \end{bmatrix} \quad AA^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

سنرى لاحقًا كيفية حساب A^{-1}، ولكن بعد أن ندرس موضوع المحددات في الوقت الحالي، لاحظ الخصائص التالية في معكوس المصفوفة :

1- $(AB)^{-1} = B^{-1}A^{-1}$ أي أن معكوس حاصل ضرب مصفوفتين هو حاصل ضرب معكوسهما بعكس الترتيب .

2- $(A^{-1})' = (A')^{-1}$، أي أن مدور معكوس المصفوفة A هو معكوس مدور المصفوفة A .

4.B. المحددات : DETERMINANTS

لكل مصفوفة مربعة A، يوجد رقم (ثابت) معروف باسم محدد المصفوفة، ويسمى det A أو يرمز له |A| حيث | | تعني "محدد خاص بـ" . لاحظ أنه لايوجد ما يسمى قيمة رقمية للمصفوفة، ولكن محدد المصفوفة عبارة عن رقم .

$$A = \begin{bmatrix} 1 & 3 & -7 \\ 2 & 5 & 0 \\ 3 & 8 & 6 \end{bmatrix} \quad |A| = \begin{vmatrix} 1 & 3 & -7 \\ 2 & 5 & 0 \\ 3 & 8 & 6 \end{vmatrix}$$

الـ |A| هو مثال لما يقال عنه محدد من الدرجة 3، حيث إنه متعلق بمصفوفة من الدرجة 3x 3 .

حساب المحدد : Evaluation of a determinant

عملية إيجاد قيمة المحدد تعرف باسم حساب أو فك أو تخفيض المحدد، ويتم ذلك من خلال التعامل مع كل عناصر المصفوفة وفقًا لشكل محدد كالتالي :

حساب قيمة محدد 2 x 2 : Evaluation of 2 x 2 Determinant

إذا كان :

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

فإن محدد هذه المصفوفة هو كالتالي :

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

وينتج ذلك من الضرب التقاطعي لعناصر القطر الرئيسي مع طرحها من الضرب التقاطعي لعناصر القطر الآخر للمصفوفة A، كما هو موضح بالأسهم.

حساب قيمة محدد 3x 3 Evaluation of a 3x 3 Determinant

إذا كان :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

وبالتالي فإن :

$$|\mathbf{A}| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{12}a_{32} - a_{13}a_{22}a_{31}$$

النظر بدقة إلى حساب محدد من الدرجة 3 x 3 يوضح التالي :

1 - كل مقدار ناتج من فك المحدد يحتوي على عنصر واحد فقط من كل صف وكل عمود.

2 - عدد العناصر في كل مقدار هي نفسها عدد الصفوف (أو الأعمدة) في المصفوفة. وبالتالي محدد 2 x 2 سيكون له عنصران في مقدار من مقادير فك هذا المحدد، وإذا كان المحدد 3 x 3 فسوف تكون هناك ثلاثة عناصر في كل مقدار من مقادير فك المحدد وهكذا.

3 - المقادير في الفك تتبادل الإشارة من + إلى -.

4 - المحدد 2 x 2 له مقداران فقط عند الفك، والمحدد 3 x 3 له ستة مقادير. القاعدة العامة: المحدد من الدرجة $N \times N$ له: $N! = N(N-1)(N-2) \ldots 3.\ 2.1$ مقدار عند الفك، حيث $N!$ تقرأ "مضروب N" وباتباع هذه القاعدة، فإن محدداً من الدرجة 5 x 5 سيكون لديه $5.\ 4.\ 3.\ 2.\ 1 = 120$ مقدار عند حساب قيمة المحدد (أو عند فكه)[1].

خصائص المحددات : Properties of determinants

1 - المصفوفة التي يكون محددها يساوي الصفر تسمى مصفوفة منفردة. في حين المصفوفة التي يكون محددها لايساوي الصفر تسمى مصفوفة غير منفردة. معكوس المصفوفة الذي ذكرناه من قبل لايتواجد في حالة المصفوفة المنفردة.

(1) الحساب قيمة محدد لمصفوفة $N \times N$ ، A ، انظر في المراجع.

2 - إذا كانت كل عناصر أي صف في **A** تساوي الصفر، فإن محدده يساوي الصفر، وبالتالي:

$$|\mathbf{A}| = \begin{vmatrix} 0 & 0 & 0 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{vmatrix} = 0$$

3 - إذا كانت $|\mathbf{A}| = |\mathbf{A}'|$ فإن ذلك يعني أن محدد **A** ومدورها متساويان.

4 - التبديل بين أي صفين أو أي عمودين متتاليين يغير إشارة $|\mathbf{A}|$.

مثال:

إذا كان

$$\mathbf{A} = \begin{bmatrix} 6 & 9 \\ -1 & 4 \end{bmatrix} \quad و \quad \mathbf{B} = \begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$$

B تم الحصول عليها بعد تبديل صفوف **A**، وبالتالي:

$|\mathbf{A}| = 24- (-9)$ و $|\mathbf{B}|= -9- (24)$

$=33$ $\quad$ $= -33$

5 - إذا تم ضرب كل عنصر من عناصر أي صف أو عمود في المصفوفة **A** في ثابت λ فإن $|\mathbf{A}|$ يكون مضروبًا أيضًا في λ.

مثال:

إذا كانت:

$$\mathbf{A} = \begin{bmatrix} 5 & -8 \\ 2 & 4 \end{bmatrix} \quad و \quad 5 = \lambda$$

وضربنا الصف الأول في 5 نحصل على:

$$\mathbf{B} = \begin{bmatrix} 25 & -40 \\ 2 & 4 \end{bmatrix}$$

سبق ورأينا أن $36 = |\mathbf{A}|$ و $180 = |\mathbf{B}|$ وذلك يساوي $5|\mathbf{A}|$.

6 - إذا كان هناك صفان أو عمودان متساويان تمامًا، فإن محدد المصفوفة يساوي الصفر.

7 - إذا كان هناك صف أو عمود في مصفوفة ما عبارة عن حاصل ضرب صف أو عمود آخر في نفس المصفوفة، فإن محددها يساوي الصفر. أي أن:

$$\mathbf{A} = \begin{bmatrix} 4 & 8 \\ 2 & 4 \end{bmatrix}$$

نجد أن عناصر الصف الأول في A عبارة عن مرتين الصف الثاني، وبالتالي 0 =|A|. بوجه عام، إذاكان أي صف (عمود) في المصفوفة عبارة عن توليفة خطية من أي صف (أو عمود) آخر، فإن المحدد= الصفر.

8 - |B| |A| =|AB| أي أن محدد حاصل ضرب مصفوفتين يساوي حاصل ضرب محدداتهما الفردية.

رتبة المصفوفة : Rank of Matrix

رتبة المصفوفة هي درجة أكبر مصفوفة جزيئية مربعة محددها لايساوي الصفر.

مثال :

$$\mathbf{A} = \begin{bmatrix} 3 & 6 & 6 \\ 0 & 4 & 5 \\ 3 & 2 & 1 \end{bmatrix}$$

يمكن إثبات أن 0 =|A|، بمعنى آخر A هي مصفوفة منفردة، وبالتالي على الرغم من أنها من الدرجة 3x 3، إلا أن رتبتها أقل من 3. في الواقع رتبة هذه المصفوفة 2، حيث إننا يمكن أن نجد مصفوفة جزيئية 2x 2 محددها لايساوي الصفر. على سبيل المثال، إذا حذفنا الصف الأول والعمود الأول من A نحصل على:

$$\mathbf{B} = \begin{bmatrix} 4 & 5 \\ 2 & 1 \end{bmatrix}$$

وهذه المصفوفة محددها 6−، وبالتالي لايساوي الصفر، ومن ثم فإن رتبة المصفوفة A هي 2. كما لاحظنا من قبل مقلوب أو معكوس المصفوفة المنفردة غير موجود، وبالتالي له مصفوفة $N \times N$ لابد أن تكون رتبتها N كي يوجد لها معكوس، إذا كان أقل من N فهي مصفوفة منفردة.

الثانوي : Minor

إذا كان الصف رقم I والعمود رقم j في مصفوفة $N \times N$ A تم حذفهما، فإن المحدد الخاص بالمصفوفة الجزئية الناتجة يسمى المحدد الثانوي للعنصر a_{ij} (عنصر التقاطع بين الصف i والعمود j) ويرمز له $|\mathbf{M}_{ij}|$.

مثال :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

المحدد الثانوي لـ a_{11} هو :

$$|\mathbf{M_{11}}| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}$$

وبالمثل المحدد الثانوي لـ a_{21} هو :

$$|\mathbf{M_{21}}| = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = a_{12}a_{33} - a_{13}a_{32}$$

المحددات الثانوية لباقي عناصر A يمكن الحصول عليها بالمثل .

المرافق : Cofactor

مرافق العنصر a_{ij} في المصفوفة $N \times N$ A يرمز له بالرمز c_{ij} يعرف كالتالي :

$$c_{ij} = (-1)^{i+j}\,|\mathbf{M}_{ij}|$$

بمعنى آخر فالمرافق هو ثانوي ذو إشارة معينة . الإشارة تكون موجبة إذا كان i+j زوجيًا ، وتكون سالبة إذا كان $i+j$ فرديًا . وبالتالي المرافق الخاص بالعنصر a_{11} في مصفوفة 3 x 3 A المعطاة سابقًا هو $a_{22}a_{33}-a_{23}a_{32}$ في حين أن مرافق العنصر a_{21} هو $-(a_{12}a_{33}-a_{13}a_{32})$ بما أن جمع 1 , 2 هو 3 وذلك رقم فردي .

مصفوفة المرافقة : Cofactor Matrix

باستبدال عناصر a_{ij} للمصفوفة A بالمرافقات الخاصة بها ، نحصل على مصفوفة معروفة باسم مصفوفة المرافقات لـ A ويرمز لها (cof A) .

المصفوفة المجاورة Adjoint Matrix : المصفوفة المجاورة والتي تكتب (adj A) عبارة عن مدور مصفوفة المرافقات ، أي أن $(\text{adj } \mathbf{A}) = (\text{cof } \mathbf{A})'$.

5.B إيجاد معكوس مصفوفة مربعة :
FINDING THE INVERSE OF A SQUARE MATRIX

إذا كانت A مصفوفة مربعة وغير منفردة (أي أن $|\mathbf{A}| \neq 0$) فإن معكوس المصفوفة $\mathbf{A}^{-1}$ يمكن إيجاده كالتالي :

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}(\text{adj } \mathbf{A})$$

الخطوات التي يتم بها إيجاد معكوس المصفوفة هي:

1 - اوجد محدد A. إذا لم يكن يساوي الصفر. اذهب إلى الخطوة 2.

2 - استبدل كل عنصر a_{ij} في A بالمرافق الخاص به للحصول على مصفوفة المرفقات.

3 - اوجد مدور مصفوفة المرافقات للحصول على المصفوفة المجاورة.

4 - اقسم لكل عنصر في المصفوفة المجاورة على |A|.

مثال:

اوجد معكوس المصفوفة

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 4 \\ 2 & 1 & 3 \end{bmatrix}$$

الخطوة 1: نحتاج أولاً حساب محدد المصفوفة. بتطبيق قواعد حساب قيمة محدد 3x 3 السابق ذكرها نحصل على 24- =|A|.

الخطوة 2: نحصل الآن على مصفوفة المرافقات، مثلاً C.

$$C = \begin{bmatrix} \begin{vmatrix} 7 & 4 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 5 & 4 \\ 2 & 3 \end{vmatrix} & \begin{vmatrix} 5 & 7 \\ 2 & 1 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 7 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 5 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 5 & 7 \end{vmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 17 & -7 & -9 \\ -3 & -3 & 3 \\ -13 & 11 & -3 \end{bmatrix}$$

الخطوة 3: ندور مصفوفة المرافقات السابقة فنحصل على المصفوفة المجاورة التالية:

$$(\text{adj } A) = \begin{bmatrix} 17 & -3 & -13 \\ -7 & -3 & 11 \\ -9 & 3 & -3 \end{bmatrix}$$

الخطوة 4: نقوم الآن بقسمة عناصر المصفوفة المجاورة على قيمة المحدد المساوية لـ 24- فنحصل على

$$A^{-1} = -\frac{1}{24} \begin{bmatrix} 17 & -3 & -13 \\ -7 & -3 & 11 \\ -9 & 3 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{17}{24} & \frac{3}{24} & \frac{13}{24} \\ \frac{7}{24} & \frac{3}{24} & -\frac{11}{24} \\ \frac{9}{24} & -\frac{3}{24} & \frac{3}{24} \end{bmatrix}$$

ومن السهل إثبات أن :

$$AA^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

وهذه مصفوفة الوحدة. يمكن للقارئ إثبات أنه بالنسبة للمثال التوضيحي المعطى في الملحق C، فإن معكوس المصفوفة X'X مساو لما هو موجود في المعادلة (5.10.C).

6.B تفاضل المصفوفات : MATRIX DIFFERENTIATION

حتى تستطيع فهم المواضيع المشروحة في ملحق CA، فقرة 2.CA نحتاج إلى بعض القواعد الخاصة بتفاضل المصفوفات.

القاعدة 1

إذا كانت $a' = [a_1\ a_2\ \dots\ a_n]$ متجه صفي من الأرقام

و

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

متجه عمودي من المتغيرات $X_1, X_2, \dots, X_n$ فإن :

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

القاعدة 2

اعتبر المصفوفة x'Ax كالتالي :

$$\mathbf{x}'\mathbf{A}\mathbf{x} = [x_1\quad x_2\quad \cdots\quad x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

وبالتالي فإن :

$$\frac{\partial(\mathbf{x}'\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}$$

وهذا متجه عمود من n عنصر،

أو
$$\frac{\partial(\mathbf{x}'\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{x}'\mathbf{A}$$

ويعتبر متجهًا صفيًا من n عنصر

REFERENCES — المراجع :

Chiang, Alpha C.: Fundamental Methods of Mathematical Economics, 3d ed., McGraw-Hill, New York, 1984, chaps. 4 and 5. This is an elementary discussion.

Hadley, G.: LinearAlgebra, Addison-Wesley, Reading, Mass., 1961. This is an advanced discussion.

ملحق C

طريقة المصفوفات لنماذج الانحدار الخطي

The Matrix Approach to Linear Regression Model

في هذا الملحق ، يتم استعراض نموذج الانحدار الخطي التقليدي الذي يوجد فيه k متغير (Y و X_2، X_3، . . . ، X_k) باستخدام رموز جبر المصفوفات . من حيث الفكرة ، فإن النموذج الذي يحتوي على k متغير هو امتداد طبيعي لنماذج المتغيرين والثلاثة متغيرات التي تم مناقشتها في هذا الكتاب . وبالتالي ، ففي هذا الملحق، يتم تقديم القليل من المفاهيم الجديدة في صورة مصفوفات[1] .

1.C نموذج الانحدار الخطي ذو K متغير :
THE K- VARIABLE LINEAR REGRESSION MODEL

إذا قمنا بتعميم نماذج الانحدار الخطي ذات المتغيرين أو الثلاثة، فنجد أن نموذج انحدار مجتمع به k متغير (PRF) فيهم المتغير التابع Y و $k-1$ متغير مفسر $X_2, X_3, ..., X_k$ تتم كتابته كالتالي :

$$\text{PRF: } Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \qquad i = 1, 2, 3, \ldots, n$$

(1.1.C)

حيث β_1 هو الجزء الثابت المقطوع من المحور الصادي، β_2 إلى β_k هي معاملات الميل الجزيئية، u هو مقدار الخطأ العشوائي و i = المفردة رقم i، n هي حجم المجتمع الـ PRF (1.1.C) يتم تفسيره بالطريقة العادية : يعطي المتوسط أو القيمة المتوسطة لـ Y مشروطة بقيم ثابتة (في عينات متكررة) لـ X_2، X_3، . . . ، X_k أي $E(Y \mid X_{2i}, X_{3i}, ..., X_{ki})$.

(1) القارئ غير المعتاد على جبر المصفوفات، عليه مراجعة ملحق B قبل قراءة المزيد من التفاصيل . ملحق B يعطي أساسيات جبر المصفوفات اللازمة لفهم هذا الملحق .

المعادلة: (1.1.C) هي اختصار للمجموعة التالية من n معادلة آنية كالتالي:

$$
\begin{aligned}
Y_1 &= \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_k X_{k1} + u_1 \\
Y_2 &= \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_k X_{k2} + u_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
Y_n &= \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_k X_{kn} + u_n
\end{aligned} \tag{2.1.C}
$$

دعنا نكتب هذا النظام من المعادلات (2.1.C) بشكل بديل، ولكنه أكثر وضوحاً كالتالي[2]:

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \tag{3.1.C}
$$

$$
\underset{n \times 1}{\mathbf{y}} = \underset{n \times k}{\mathbf{X}} \; \underset{k \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\mathbf{u}}
$$

حيث إن: $\mathbf{y}$ = متجه عمودي $n \times 1$ من مفردات المتغير التابع Y

$\mathbf{X}$ = مصفوفة $n \times k$ تعطي n مفردة خاصة بـ $k-1$ متغير X_2 إلى X_k العمود الأول يمثل أولاً المقدار الثابت (هذه المصفوفة معروفة أيضاً باسم مصفوفة البيانات).

$\boldsymbol{\beta}$ = متجه عمودي $k \times 1$ من المعلمات المجهولة β_1، β_2، ...، β_k.

$\mathbf{u}$ = متجه عمودي $n \times 1$ من مقادير الأخطاء u_i.

باستخدام قواعد ضرب المصفوفات وجمعها، يمكن للقارئ أن يتأكد من أن النظام (2.1.C) و (3.1.C) متكافئان.

النظام (3.1.C) معروف باسم صياغة نموذج الانحدار الخطي العام (k متغير) في صورة مصفوفات. ويمكن كتابته باختصار كالتالي:

$$
\underset{n \times 1}{\mathbf{y}} = \underset{n \times k}{\mathbf{X}} \; \underset{k \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\mathbf{u}} \tag{4.1.C}
$$

حيث لايوجد أي لبس في أبعاد أو درجات المصفوفة X والمتجهات $\mathbf{y}$، $\boldsymbol{\beta}$ و $\mathbf{u}$ المعادلة (4.1.C) يمكن كتابتها ببساطة كالتالي:

$$
\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{5.1.C}
$$

(2) Following the notation introduced in **App. B**, we shall represent vectors by lowercase bold-faced letters and matrices by uppercase boldfaced letters.

لشرح الصياغة بالمصفوفات، اعتبر نموذج الدخل- الاستهلاك ذو المتغيرين والذي درسناه في الفصل (3) وهو $Y_i = \beta_1 + \beta_2 X_i + u_i$ حيث y هي نفقات الاستهلاك وX هو الدخل. باستخدام البيانات المعطاة في جدول (3.2). يمكن كتابة معادلة المصفوفة السابقة كالتالي:

$$\begin{bmatrix} 70 \\ 65 \\ 90 \\ 95 \\ 110 \\ 115 \\ 120 \\ 140 \\ 155 \\ 150 \end{bmatrix} = \begin{bmatrix} 1 & 80 \\ 1 & 100 \\ 1 & 120 \\ 1 & 140 \\ 1 & 160 \\ 1 & 180 \\ 1 & 200 \\ 1 & 220 \\ 1 & 240 \\ 1 & 260 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \\ u_{10} \end{bmatrix} \qquad (6.1.C)$$

$$\underset{10 \times 1}{\mathbf{y}} = \underset{10 \times 2}{\mathbf{X}} \; \underset{2 \times 1}{\boldsymbol{\beta}} + \underset{10 \times 1}{\mathbf{u}}$$

كما في حالة المتغيرات الثنائية والثلاثية، فإن هدفنا هو تقدير معالم الانحدار المتعددة (1.1.C) وعمل استدلال إحصائي حولهما من البيانات المتاحة. في رموز المصفوفة نريد تقدير $\boldsymbol{\beta}$ واستنتاج استدلالات إحصائية حول $\boldsymbol{\beta}$. بفرض التقدير، يمكن أن تستخدم طريقة المربعات الصغرى العادية (OLS) أو طريقة الإمكان الأعظم (ML). ولكن كما لاحظنا من قبل، فإن هاتين الطريقتين تعطيان تقديرات متساوية تماماً لمعاملات الانحدار[3] وبالتالي سنهتم بطريقة الـ OLS.

2.C فروض نموذج الانحدار الخطي التقليدي في صورة مصفوفات: ASSUMPTIONS OF THE CLASSICAL LINEAR REGRESSION MODEL IN MATRIX NOTATION

الفروض الخاصة لنموذج الانحدار التقليدي المعطاة في جدول (1.C)، يتم استعراضها بطريقة رموز فردية ثابتة وبرموز المصفوفات أيضاً. الفرض 1 المعطى في (1.2.C) يعني أن القيمة المتوقعة لمقدار الخطأ $\mathbf{u}$، أي لكل عنصر من عناصره، هو الصفر. بشكل أكثر دقة، فإن $E(\mathbf{u}) = \mathbf{0}$ تعني أن:

---

(3) لإثبات ذلك في حالة وجود k متغير، ارجع إلى المراجع المذكورة في ملاحظات الفصل (4).

$$E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} E(u_1) \\ E(u_2) \\ \vdots \\ E(u_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \textbf{(1.2.C)}$$

جدول (1.C) فروض نموذج الانحدار الخطي التقليدي

الرموز التقليدية		رموز المصفوفات
1- لكل i $\quad E(u_i) = 0$	(1.2.3)	1 - $E(\mathbf{u})= 0$ حيث **u** و **0** هما متجهان عموديان nx1 و **0** هو المتجه الصفري
2 - $i \neq j$ $\quad E(u_i u_j) = 0$ $i = j$ $\quad =\sigma^2$	(5.2.3)	2 - $E(\mathbf{uu}')= \sigma^2\mathbf{I}$ حيث **I** هي مصفوفة الوحدة n x n
3 - $X_2, X_3, ..., X_k$ متغيرات غير عشوائية أو محددة	(2.2.3)	3 - المصفوفة **X** (n x k) هي مصفوفة غير عشوائية أي إنها تتكون من مجموعة من الأرقام الثابتة المحددة .
4 - لاتوجد علاقة خطية محددة بين متغيرات الـ X أي أنه لايوجد ارتباط متعدد	(7.1.7)	4 - رتبة المصفوفة **X** هي $p(\mathbf{X})= k$ حيث k هو عدد أعمدة **X** و k أقل من عدد المشاهدات n .
5 -لاختبارات الفروض فإن $u_j \sim N(0, \sigma^2)$	(4.2.4)	5 - المتجه **u** له التوزيع الطبيعي المتعدد أي أن $\mathbf{u}\sim N(\mathbf{0}, \sigma^2\mathbf{I})$

الفرض 2 [المعادلة (2.2.C)] هو طريقة مدمجة للتعبير عن الفرضين الموجودين في (5.2.3) و (2.2.3) بالرموز التقليدية . ليتضح ذلك دعنا نكتب :

$$E(\mathbf{uu}') = E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} [u_1 \quad u_2 \quad \cdots \quad u_n]$$

حيث **u′** هو مدور المتجه العمودي **u** وهو متجه صفي . وبالقيام بالضرب نحصل على :

$$E(\mathbf{uu}') = E\begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & \cdots & u_2u_n \\ \cdots & \cdots & \cdots & \cdots \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix}$$

وبإدخال علامة التوقع E على كل عنصر في المصفوفة السابقة، فإننا نحصل على:

$$E(\mathbf{uu}') = \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \cdots & \cdots & \cdots & \cdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{bmatrix} \qquad (2.2.C)$$

وحيث إن هناك فرض ثبات التباين، وعدم وجود ارتباط تسلسلي، فإن المصفوفة (2.2.C) تكتب على الشكل التالي:

$$E(\mathbf{uu}') = \begin{bmatrix} \sigma^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (3.2.C)$$

$$= \sigma^2 \mathbf{I}$$

حيث I هي مصفوفة الوحدة $n \times n$.

المصفوفة (2.2.C) [والصورة الموجودة عليها في (3.2.C)] تسمى مصفوفة التباين- التغاير لمقادير الأخطاء u_i، العناصر الموجودة على القطر الرئيسي لهذه المصفوفة (بداية من أعلى الشمال إلى الركن الأسفل على اليمين) تمثل التباين والعناصر الموجودة على جانبي القطر الرئيسي تمثل التغاير[4]. لاحظ أن مصفوفة التباين- التغاير مصفوفة متماثلة، العناصر أعلى القطر الرئيسي وأسفله متساويان.

الفرض 3 يعني أن المصفوفة $n \times k$ ، $\mathbf{X}$ هي مصفوفة غير عشوائية، أي أنها تتكون من أرقام ثابتة، كما لاحظنا من قبل، فإن تحليل الانحدار هو تحليل انحدار شرطي، حيث إنه مشروط على قيم ثابتة للمتغيرات X.

الفرض 4 يعني أن المصفوفة $\mathbf{X}$ لها رتبة كاملة مساوية لـ k، حيث k هو عدد أعمدة المصفوفة. أي أن ذلك يعني أن أعمدة المصفوفة X مستقلة خطياً. بمعنى آخر، لا يوجد ارتباط متعدد. في الترميز التقليدي (الثوابت) فإن ذلك

(4) بالتعريف، تباين u_i هو $u_i = E[u_i - E(u_i)]^2$ والتغاير بين u_i و u_j هو $E[u_i - E(u_i)][u_j - E(u_i)]$ ولكن بسبب الفرض $E(u_i) = 0$ لكل i فيكون لدينا مصفوفة التباين والتغاير الموجودة في (3.2.C).

مساوي للقول بأنه لاتوجد مجموعة أخرى من الأرقام لـ λ_1، λ_2، . . . ، λ_k ليست جميعاً تساوي الصفر حيث [CF. (7.1.8)] .

$$\lambda_1 X_{1i} + \lambda_2 X_{2i} + \dots + \lambda_k X_{ki} = 0 \qquad (4.2.C)$$

حيث : $X_{1i} = 1$ لكل i (حتى يكون هناك عمود من 1's في المصفوفة **X** .

باستخدام المصفوفات فإن : (4.2.C) يمكن كتابتها كالتالي :

$$\lambda' \mathbf{x} = 0 \qquad (5.2.C)$$

حيث : λ' هو متجه صفي $1 \times k$ و **x** هو متجه عمودي $k \times 1$.

إذا وجدت علاقة خطية تامة كالموجودة في (4.2.C) فإن المتغيرات يقال عنها مرتبطة خطياً ، في حين إذا تحققت (4.2.C) فقط عندما يكون $\lambda_1 = \lambda_2 = \lambda_3 = \dots = 0$ فإن المتغيرات X يقال عنها مستقلة خطياً. لأسباب أكثر وضوحاً لفرض عدم وجود ارتباط خطي متعدد ، ارجع إلى الفصل (7) وأيضاً تم استعراض هذا الفرض بعمق في الفصل (10) .

3.C تقدير OLS : OLS ESTIMATION

للحصول على تقدير OLS لـ $\boldsymbol{\beta}$، دعنا أولاً نكتب انحدار عينة k من المتغيرات (SRF) كالتالي :

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + \hat{u}_i \qquad (1.3.C)$$

والذي يمكن كتابته باختصار وباستخدام المصفوفات كالتالي :

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} \qquad (2.3.C)$$

باستخدام المصفوفات

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{bmatrix} \qquad (3.3.C)$$

$$\underset{n \times 1}{\mathbf{y}} = \underset{n \times k}{\mathbf{X}} \quad \underset{k \times 1}{\hat{\boldsymbol{\beta}}} + \underset{n \times 1}{\hat{\mathbf{u}}}$$

حيث $\hat{\boldsymbol{\beta}}$ هو متجه عمودي به k عنصر لتقدير OLS لمعاملات الانحدار و $\hat{\mathbf{u}}$ هو متجه عمودي $n \times 1$ مكون من n من البواقي .

كما في حالة النماذج ذات المتغيرين أو الثلاثة متغيرات ، فإن مقدرات OLS في حالة وجود k متغير يمكن الحصول عليها بتصغير التالي :

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})^2 \qquad (4.3.C)$$

حيث $\Sigma \hat{\mathbf{u}}_i^2$ هو مجموع مربعات البواقي (RSS). باستخدام المصفوفات فإننا نريد تصغير المقدار $\hat{\mathbf{u}}'\hat{\mathbf{u}}$ حيث إن:

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = [\hat{u}_1 \quad \hat{u}_2 \quad \cdots \quad \hat{u}_n] \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{bmatrix} = \hat{u}_1^2 + \hat{u}_2^2 + \cdots + \hat{u}_n^2 = \sum \hat{u}_i^2 \qquad (5.3.C)$$

والآن من (2.3.C) نحصل على:

$$\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} \qquad (6.3.C)$$

وبالتالي:

$$\begin{aligned} \hat{\mathbf{u}}'\hat{\mathbf{u}} &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned} \qquad (7.3.C)$$

حيث تم استخدام خاصية مدور المصفوفة وهي $(\mathbf{X}\ \hat{\boldsymbol{\beta}})' = \hat{\boldsymbol{\beta}}'\mathbf{X}'$، وبما أن $\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$ ثابت (رقم حقيقي) فإنه يساوي مدورة $\mathbf{y}'\mathbf{X}\ \hat{\boldsymbol{\beta}}$.

المعادلة (7.3.C) تعتبر نفس (4.3.C) ولكن في صورة مصفوفات، في الرموز التقليدية. فإن طريقة OLS تهدف إلى تقدير β_1، β_2، . . . ، β_k، حيث إن $\sum \hat{u}_i^2$ يكون أصغر ما يمكن. ويتم الوصول إلى ذلك بتفاضل (4.3.C) جزئياً بالنسبة إلى $\hat{\beta}_1$، $\hat{\beta}_2$، . . . ، $\hat{\beta}_k$ وتساوي الناتج في كل مرة بالصفر. وهذه العملية تؤدي إلى k من المعادلات الآنية في k من المجاهيل وهذا تبعاً لنظرية المربعات الصغرى للمعادلات الطبيعية. وكما رأينا في ملحق A، فقرة 1.CA فإن هذه المعادلات كالتالي:

$$\begin{aligned} n\hat{\beta}_1 + \hat{\beta}_2 \sum X_{2i} + \hat{\beta}_3 \sum X_{3i} + \cdots + \hat{\beta}_k \sum X_{ki} &= \sum Y_i \\ \hat{\beta}_1 \sum X_{2i} + \hat{\beta}_2 \sum X_{2i}^2 + \hat{\beta}_3 \sum X_{2i}X_{3i} + \cdots + \hat{\beta}_k \sum X_{2i}X_{ki} &= \sum X_{2i}Y_i \\ \hat{\beta}_1 \sum X_{3i} + \hat{\beta}_2 \sum X_{3i}X_{2i} + \hat{\beta}_3 \sum X_{3i}^2 + \cdots + \hat{\beta}_k \sum X_{3i}X_{ki} &= \sum X_{3i}Y_i \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{\beta}_1 \sum X_{ki} + \hat{\beta}_2 \sum X_{ki}X_{2i} + \hat{\beta}_3 \sum X_{ki}X_{3i} + \cdots + \hat{\beta}_k \sum X_{ki}^2 &= \sum X_{ki}Y_i \end{aligned}$$

(C.3.8)[5]

(5) هذه المعادلات يمكن تذكرها بسهولة. ابدأ بالمعادلة $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki}$ وبجمع هذه المعادلة على الـ n قيمة نحصل على المعادلة الأولى في (8.3.C)، ثم نضربها في X_2 من الناحيتين ونجمع على n مرة أخرى فنحصل على المعادلة الثانية ثم نضرب بعد ذلك في X_3 من الناحيتين ونجمع على n فنحصل على المعادلة الثالثة وهكذا. وعموماً لاحظ أن المعادلة الأولى في (8.3.C) تعطي مباشراً $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}_2 - \cdots - \hat{\beta}_k \bar{X}_k$ [cf. (7.4.6)]

$$\begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} & \cdots & \sum X_{ki} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} & \cdots & \sum X_{2i}X_{ki} \\ \sum X_{3i} & \sum X_{3i}X_{2i} & \sum X_{3i}^2 & \cdots & \sum X_{3i}X_{ki} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum X_{ki} & \sum X_{ki}X_{2i} & \sum X_{ki}X_{3i} & \cdots & \sum X_{ki}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & \cdots & X_{2n} \\ X_{31} & X_{32} & \cdots & X_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix}$$

$$\quad (\mathbf{X'X}) \qquad\qquad \hat{\boldsymbol{\beta}} \qquad\qquad \mathbf{X'} \qquad\qquad \mathbf{y} \qquad (9.3.C)$$

$$(\mathbf{X'X})\hat{\boldsymbol{\beta}} = \mathbf{X'y} \qquad (10.3.C)$$

لاحظ الخصائص التالية للمصفوفة $(\mathbf{X'X})$: (1) أنها تعطي مجموع مربعات وحواصل الضرب التبادلية للمتغيرات X، واحد منها هو المقدار الثابت المقطوع من المحور الصادي، وذلك بوضع القيمة 1 لكل مشاهدة. العناصر الموجودة على القطر الرئيسي تعطي مجموع المربعات الطبيعي، والعناصر الموجودة على جانبي القطر الرئيسي تعطي مجموع حواصل الضرب التبادلية الطبيعية (المقصود بالطبيعي أي أن الوحدات الأصلية مقاسة بالمقياس الأصلي لها). (2) هذه المصفوفة تعتبر مصفوفة متماثلة، حيث إن الضرب التبادلي بين X_{2i} و X_{3i} هو نفسه بين X_{3i} و X_{2i}. (3) إنها مصفوفة من الدرجة $(k \times k)$ أي أن لها k صف و k عمود.

في (10.3.C) الكميات المعلومة هي $(\mathbf{X'X})$ و $(\mathbf{X'y})$ (الضرب التبادلي بين متغيرات X و y) والمجهول هو $\hat{\boldsymbol{\beta}}$. وباستخدام جبر المصفوفات، فإن معكوس $(\mathbf{X'X})$ يمكن إيجاده وهو $(\mathbf{X'X})^{-1}$ وبضرب كل من طرفي (10.3.C) في هذا المعكوس نحصل على:

$$(\mathbf{X'X})^{-1}(\mathbf{X'X})\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$$

ولكن حيث إن: $(\mathbf{X'X})^{-1}(\mathbf{X'X}) = \mathbf{I}$، مصفوفة من الدرجة $k \times k$، فإن:

$$\mathbf{I}\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$$

أو

$$\underset{k \times 1}{\hat{\boldsymbol{\beta}}} = \underset{k \times k}{(\mathbf{X'X})^{-1}} \; \underset{(k \times n)}{\mathbf{X'}} \; \underset{(n \times 1)}{\mathbf{y}} \qquad (11.3.C)$$

المعادلة (11.3.C) هي نتيجة أساسية في نظرية OLS باستخدام المصفوفات، حيث توضح أن المتجه $\hat{\boldsymbol{\beta}}$ يمكن تقديره من البيانات المعطاة، فعلى الرغم من أن (11.3.C) تم

الحصول عليه من (9.3.C)، فإنه يمكن الحصول عليه مباشرة من (7.3.C) بتفاضل $\mathbf{\hat{u}'\hat{u}}$ بالنسبة لـ $\hat{\beta}$. الإثبات معطى في ملحق CA، الفقرة 2.CA.

مثال توضيحي : An Illustration Example

لتوضيح طريقة المصفوفات التي تم استعراضها حتى الآن، دعنا نعيد استخدام مثال الدخل - الاستهلاك الموجود في الفصل (3)، وبياناته معطاة في (6.1.C) في حالة وجود متغيرين اثنين يكون لدينا:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$(\mathbf{X'X}) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_1 & X_2 & X_3 & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ & \cdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

و

$$\mathbf{X'y} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_1 & X_2 & X_3 & \cdots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

وباستخدام البيانات المعطاة في (6.1.C) نحصل على:

$$\mathbf{X'X} = \begin{bmatrix} 10 & 1700 \\ 1700 & 322000 \end{bmatrix}$$

و

$$\mathbf{X'y} = \begin{bmatrix} 1110 \\ 205500 \end{bmatrix}$$

وباستخدام قواعد المصفوفات للحصول على معكوس المصفوفة المعطاة في الملحق B، يمكن أن نثبت أن معكوس المصفوفة $(\mathbf{X'X})$ السابقة هو:

$$\mathbf{X'X}^{-1} = \begin{bmatrix} 0.97576 & -0.005152 \\ -0.005152 & 0.0000303 \end{bmatrix}$$

وبالتالي :

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 0.97576 & -0.005152 \\ -0.005152 & 0.0000303 \end{bmatrix} \begin{bmatrix} 1110 \\ 205500 \end{bmatrix}$$

$$= \begin{bmatrix} 24.4545 \\ 0.5079 \end{bmatrix}$$

من قبل حصلنا على $\hat{\beta}_1 = 24.4541$ و $\hat{\beta}_2 = 0.5091$ باستخدام الحاسب الآلي. الفرق بين الطريقتين يرجع إلى خطأ التقريب. عموماً لاحظ أن استخدام آلة حاسبة ضروري للحصول على نتائج بأرقام كسرية عددية لتقليل أخطاء التقريب.

مصفوفة التباين - التغاير لـ $\hat{\boldsymbol{\beta}}$

طرق المصفوفات تؤهلنا لعمل معادلات ليس فقط للحصول على تباين $\hat{\beta}_i$، أي عنصر في $\hat{\boldsymbol{\beta}}$، ولكن أيضاً التغاير يبين أي عنصرين في $\hat{\boldsymbol{\beta}}$، مثلاً $\hat{\beta}_i$ و $\hat{\beta}_j$. نحتاج لمثل هذه التغايرات والتباينات لفرض الاستدلال الإحصائي.

بالتعريف، فإن مصفوفة التباين والتغاير لـ $\hat{\boldsymbol{\beta}}$ هي [Cf. (2.2.C)]

$$\text{var-cov}(\hat{\boldsymbol{\beta}}) = E\{[\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}})][\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}})]'\}$$

والتي يمكن كتابتها بالتفصيل كالتالي:

$$\text{var-cov}(\hat{\boldsymbol{\beta}}) = \begin{bmatrix} \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \cdots & \text{cov}(\hat{\beta}_1, \hat{\beta}_k) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{var}(\hat{\beta}_2) & \cdots & \text{cov}(\hat{\beta}_2, \hat{\beta}_k) \\ \cdots & \cdots & \cdots & \cdots \\ \text{cov}(\hat{\beta}_k, \hat{\beta}_1) & \text{cov}(\hat{\beta}_k, \hat{\beta}_2) & \cdots & \text{var}(\hat{\beta}_k) \end{bmatrix} \quad (12.3.C)$$

في الملحق CA، فقرة 3.CA تم إثبات أن مصفوفة التباين - التغاير السابقة - يمكن الحصول عليها من المعادلة التالية:

$$\text{var-cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \quad (13.3.C)$$

حيث σ^2 هو تباين ثابت لـ u_i و $(\mathbf{X}'\mathbf{X})^{-1}$ هو معكوس المصفوفة الموجودة في المعادلة (11.3.C) والتي تعطي مقدرات OLS لـ $\hat{\boldsymbol{\beta}}$.

في نماذج الانحدار الخطية ذات المتغيرين والثلاثة متغيرات، يعتبر σ^2 مقدر غير متحيز لـ σ^2 وهو يساوي $\hat{\sigma}^2 = \sum \hat{u}_i^2/(n-2)$ ويساوي $\sum \hat{u}_i^2/(n-3)$ بالترتيب. في حالة وجود k متغير، فإن المعادلة هي:

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-k} = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k} \quad (14.3.C)$$

حيث يوجد الآن درجات حرية $n-k$. (لماذا؟).

على الرغم من أن $\hat{\mathbf{u}}'\hat{\mathbf{u}}$ يمكن حسابها من البواقي المقدرة، فإنه عملياً يمكن أن نحصل عليه مباشراً كالتالي. وتذكر أن $\sum \hat{u}_i^2 (= \text{RSS}) = \text{Tss} - \text{Ess}$، فإنه في حالة متغيرين يمكن أن يكتب كالتالي:

$$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2^2 \sum x_i^2 \tag{6.3.3}$$

وفي حالة وجود ثلاثة متغيرات فإن:

$$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2 \sum y_i x_{2i} - \hat{\beta}_3 \sum y_i x_{3i} \tag{19.4.7}$$

وبتعميم ذلك، يمكن أن نرى أن نموذج الـ k متغير يمكن كتابته كالتالي:

$$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2 \sum y_i x_{2i} - \cdots - \hat{\beta}_k \sum y_i x_{ki} \tag{15.3.C}$$

وباستخدام الرموز فإن:

$$\text{TSS: } \sum y_i^2 = \mathbf{y}'\mathbf{y} - n\bar{Y}^2 \tag{16.3.C}$$

$$\text{ESS: } \hat{\beta}_2 \sum y_i x_{2i} + \cdots + \hat{\beta}_k \sum y_i x_{ki} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2 \tag{17.3.C}$$

حيث المقدار $n\bar{Y}^2$

معروف باسم المتوسط المصحح[6]. وبالتالي:

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \tag{18.3.C}$$

بمجرد الحصول على $\hat{\mathbf{u}}'\hat{\mathbf{u}}$، فإن $\hat{\sigma}^2$ يمكن بسهولة حسابه من (14.3.C) والذي بدوره يجعلنا نستطيع تقدير مصفوفة التباين والتغاير (13.3.C).

$$\begin{aligned}\hat{\mathbf{u}}'\hat{\mathbf{u}} &= 132100 - [24.4545 \quad 0.5091]\begin{bmatrix}1110\\205500\end{bmatrix}\\ &= 337.373\end{aligned}$$

وبالتالي: $\hat{\sigma}^2 = (337.273/8) = 42.1591$ والذي يساوي تقريباً القيمة التي حصلنا عليها في الفصل (3).

(6) لاحظ أن: $\sum y_i^2 = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2 = \mathbf{y}'\mathbf{y} - n\bar{Y}^2$ وبالتالي بدون مقدار التصحيح فإن $\mathbf{y}'\mathbf{y}$ سيعطي ببساطة مجموع المربعات الصافي، وليس مجموع مربعات الانحرافات.

خصائص متجه OLS لـ $\hat{\beta}$: **Properties of OLS Vector $\hat{\beta}$**

في حالة وجود متغيرين أو ثلاثة متغيرات، نعرف أن مقدرات OLS هي مقدرات خطية وغير متحيزة وداخل فئة كل المقدرات الخطية غير المتحيزة، فإن مقدرات OLS لها أقل تباين (خاصية Gouss- Markov). باختصار، فإن مقدرات OLS هي أفضل المقدرات الخطية غير المتحيزة (BLUE). هذه الخاصية يمكن تعميمها لمتجه $\hat{\beta}$ بأكمله، أي أن $\hat{\beta}$ خطي (كل عنصر داخله هو دالة خطية في المتغير التابع Y). $E(\hat{\beta}) = \hat{\beta}$. أي أن القيمة المتوقعة لكل عنصر في $\hat{\beta}$ يساوي العنصر المناظر للقيمة الحقيقية لـ β وداخل فئة كل المقدرات الخطية غير المتحيزة فإن مقدر OLS لـ $\hat{\beta}$ له أقل تباين.

الإثبات معطى في ملحق AC، فقرة AC.4. وكما سبق وذكرنا في المقدمة، فإنه في حالة k متغير في أغلب الأحيان ما هي إلا تعميم مباشر لحالة متغيرين أو ثلاثة متغيرات.

4.C معامل التحديد R^2 باستخدام المصفوفات :
THE COEFFICIENT OF DETERMINATION R^2 IN MATRIX NOTATION

تم تعريف معامل التحديد R^2 من قبل كالتالي:

$$R^2 = \frac{\text{ESS}}{\text{TSS}}$$

في حالة متغيرين اثنين،

$$R^2 = \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum y_i^2} \qquad (6.5.3)$$

في حالة وجود ثلاثة متغيرات، فإن:

$$R^2 = \frac{\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i}}{\sum y_i^2} \qquad (5.5.7)$$

وعند التعميم في حالة k متغير، فإن:

$$R^2 = \frac{\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i} + \cdots + \hat{\beta}_k \sum y_i x_{ki}}{\sum y_i^2} \qquad (1.4.C)$$

باستخدام (16.3.C) و (17.3.C)، معادلة (1.4.C) يمكن أن تكتب كالتالي:

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2} \qquad (2.4.C)$$

والذي يمثل R^2 في صورة المصفوفات (باستخدام رموز المصفوفات) كمثال توضيحي فإن:

$$\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = [24.3571 \quad 0.5079]\begin{bmatrix} 1{,}110 \\ 205{,}500 \end{bmatrix}$$
$$= 131{,}409.831$$
$$\mathbf{y}'\mathbf{y} = 132{,}100$$

و

$$n\bar{Y}^2 = 123{,}210$$

وبالتعويض عن هذه القيم في (2.4.C) نرى أن $R^2 = 0.9224$ والذي يساوي تقريباً القيمة التي حصلنا عليها من قبل، مع وضع خطأ التقريب في الاعتبار.

5.C مصفوفة الارتباط: THE CORRELATION MATRIX

في الفصول السابقة، استعرضنا معاملات الارتباط البسيطة r_{12}، r_{13}، r_{23}، ومعاملات الارتباط الجزئية، أو معاملات الارتباط من الدرجة الأولى $r_{12.3}$، $r_{13.2}$، $r_{23.1}$ وعلاقاتها التبادلية. في حالة k متغير، سيكون لدينا ككل $k(k-1)/2$ معاملات ارتباط. (لماذا؟)، المعاملات $k(k-1)/2$ يمكن كتابتها في شكل مصفوفة تسمى **مصفوفة الارتباط R** كالتالي:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1k} \\ r_{21} & r_{22} & r_{23} & \cdots & r_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{k1} & r_{k2} & r_{k3} & \cdots & r_{kk} \end{bmatrix} = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1k} \\ r_{21} & 1 & r_{23} & \cdots & r_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{k1} & r_{k2} & r_{k3} & \cdots & 1 \end{bmatrix} \qquad (1.5.C)$$

حيث إن الترميز 1، كما سبق، يرمز إلى المتغير التابع Y (r_{12} تعني معامل الارتباط بين Y و X_2، وهكذا) وتم استخدام حقيقة أن معامل ارتباط متغير ما مع نفسه هو دائماً 1 ($r_{11} = r_{22} = r_{33} = \ldots = r_{kk} = 1$).

من مصفوفة الارتباط **R** يمكن الحصول على معاملات الارتباط من الدرجة الأولى (انظر الفصل 7)، والمعاملات من درجات أعلى $r_{12.34\ldots k}$ (انظر تمرين 4.C). العديد من برامج الحاسب الآلي تحسب المصفوفة **R**. وقد استخدمنا مصفوفة الارتباط من قبل في الفصل (10).

6.C اختبارات الفروض لمعاملات الانحدار الفردية باستخدام المصفوفات : HYPOTHESIS TESTING ABOUT INDIVIDUAL REGRESSION COEFFICIENTS IN MATRIX NOTATION

لأسباب تم استعراضها في الفصول السابقة، إذا كان هدفنا هو الاستدلال الإحصائي بالإضافة للتقدير، فإننا سنفترض أن مقدار الخطأ u_i سيتبع بعض التوزيعات الاحتمالية المعروفة.

ولأسباب أيضاً تم شرحها سابقاً، فإن في تحليل الانحدار عادة نفترض أن كل u_i تتبع التوزيع الطبيعي بتوقع يساوي الصفر وتباين ثابت σ^2 باستخدام المصفوفات لدينا التالي:

$$\mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \qquad \textbf{(1.6.C)}$$

حيث **u** و **0** هما متجهان عموديان $n \times 1$ و **I** هي مصفوفة الوحدة $n \times n$ و **0** هو المتجه الصفري.

بافتراض التوزيع الطبيعي، فإننا نعرف أنه في حالة نماذج الانحدار الخطية ذات المتغيرين والثلاثة متغيرات يحدث التالي (1) مقدرات OLS لـ $\hat{\beta}_i$ ومقدرات ML لـ $\hat{\beta}_i$ متساويان. ولكن مقدرات ML لـ σ^2 متحيزة على الرغم من أن هذا التحيز يمكن إزالته باستخدام مقدر OLS غير المتحيز σ^2، و (2) مقدر OLS $\hat{\beta}_i$ يتبع أيضاً التوزيع الطبيعي، وعموماً في حالة k متغير فإنه يمكن إثبات أن:

$$\hat{\boldsymbol{\beta}} \sim N[\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}] \qquad \textbf{(2.6.C)}$$

أي أن كل عنصر في $\hat{\boldsymbol{\beta}}$ له التوزيع الطبيعي بتوقع يساوي القيمة الحقيقية المناظرة $\boldsymbol{\beta}$ وتباينه يساوي σ^2 مضروبة في عنصر القطر الرئيسي المناسب الموجود في معكوس المصفوفة $(\mathbf{X}'\mathbf{X})^{-1}$.

وبما أنه عملياً σ^2 غير معلومة، فإنه يمكن تقديرها بـ$\hat{\sigma}^2$. وتنتقل إلى استخدام توزيع t، حيث يتبع كل عنصر في $\hat{\boldsymbol{\beta}}$ توزيع t بدرجات حرية $(n-k)$، وباستخدام الرموز، فإن ذلك يعني التالي:

$$t = \frac{\hat{\beta}_i - \beta_i}{\text{se}\,(\hat{\beta}_i)} \qquad \textbf{(3.6.C)}$$

بدرجات حرية $(n-k)$ و $\hat{\beta}_i$ هي أي عنصر في $\hat{\boldsymbol{\beta}}$.

توزيع t يمكن استخدامه بعد ذلك في اختبارات الفروض حول القيمة الحقيقية β_i ، كما يمكن استخدامه في تكوين فترات الثقة له. الطريقة الأصلية لذلك تم استعراضها سابقاً في الفصلين (5 و 8). لمثال كامل على ذلك، انظر الفقرة 10.C .

7.C اختبار معنوية الانحدار ككل: تحليل التباين باستخدام المصفوفات : TESTING THE OVERALL SIGNIFICANCE OF REGRESSION: ANALYSIS OF VARIANCE IN MATRIX NOTATION

في الفصل 8 استعرضنا أسلوب (1) ANOVA لاختبار معنوية الانحدار المقدر، أي لاختبار الفرض العدمي القائل بأن معاملات الميل (الجزيئية) الحقيقية تساوي الصفر آنياً. و(2) لتحديد دور المتغير المفسر في تفسير الـ y. أسلوب الـ ANOVA يمكن بسهولة تعميمه لحالة k متغير. تذكر أن أسلوب ANOVA يعتمد على تكوين TSS من جزيئين: الـ ESS وRSS. للتعبير عن الثلاثة مجاميع للمربعات في صورة مصفوفات معطى في (16.3.C)، (17.3.C) و (18.3.C) بالترتيب.

درجات الحرية المرتبطة بمجاميع المربعات هي $n-1$، $k-1$ و $n-k$ بالترتيب (لماذا؟) ثم وفقاً للفصل 8، جدول (1.8)، يمكنا تكوين جدول (2.C).

بافتراض أن مقادير الأخطاء u_i تتبع التوزيع الطبيعي، وأن الفرض العدمي هو $\beta_2 = \beta_3 = \ldots = \beta_k = 0$ فإنه وفقاً للفصل (8)، يمكن إثبات أن:

$$F = \frac{(\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2)/(k-1)}{(\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y})/(n-k)} \qquad (1.7.C)$$

يتبع توزيع F بدرجات حرية $k-1$ و $n-k$.

في الفصل (8)، رأينا أنه، وفقاً للفروض التي سبق وذكرناها، هناك علاقة وثيقة بين F و R^2 هي:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \qquad (11.5.8)$$

وبالتالي جدول الـ ANOVA (2.C) يمكن التعبير عنه في جدول (3.C). أحد مميزات جدول (3.C) عن جدول (2.C) أن التحليل بأكمله يمكن عمله من خلال R^2 بدون استخدام المقدار $(\mathbf{y}'\mathbf{y} - n\bar{Y}^2)$ مما يعني عدم استخدام نسبة الـ F.

جدول (2.C) جدول الـ ANOVA لنموذج الانحدار الخطي لـ k متغير باستخدام المصفوفات

مصدر التباين	SS	df	MSS
بسبب الانحدار (أي بسبب X_2، X_3، ...، X_k)	$\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2$	$k-1$	$\dfrac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{k-1}$
بسبب البواقي	$\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$	$n-k$	$\dfrac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k}$
الإجمالي	$\mathbf{y}'\mathbf{y} - n\bar{Y}^2$	$n-1$	

جدول (3.C) جدول ANOVA لـ k متغير باستخدام المصفوفات و R^2

مصدر التباين	SS	df	MSS
بسبب الانحدار (أي بسبب X_2، X_3، ...، X_k)	$R^2(\mathbf{y}'\mathbf{y} - n\bar{Y}^2)$	$k-1$	$\dfrac{R^2(\mathbf{y}'\mathbf{y} - n\bar{Y}^2)}{k-1}$
بسبب البواقي	$(1-R^2)(\mathbf{y}'\mathbf{y} - n\bar{Y}^2)$	$n-k$	$\dfrac{(1-R^2)(\mathbf{y}'\mathbf{y} - n\bar{Y}^2)}{n-k}$
الإجمالي	$\mathbf{y}'\mathbf{y} - n\bar{Y}^2$	$n-1$	

8.C اختبار قيود الخطية: اختبار F العام باستخدام المصفوفات :
TESTING LINEAR RESTRICTIONS: GENERAL F TESTING USING MATRIX NOTATION

في الفقرة 9.8 ناقشنا، باستخدام الرموز التقليدية، كيف يمكن أن تستخدم الانحدار المتعدد المقدر في التنبؤ (1) المتوسط (2) قيم y الفردية بمعلومية قيم المتغيرات المنحدرة X. في هذه الفقرة، سنرى كيف يمكن التعبير عن ذلك كله باستخدام المصفوفات. وسنستعرض أيضاً معادلات تقدير التباين والأخطاء القياسية للقيم المقدرة، في الفصل (8) لاحظنا أن هذه المعادلات تكون صياغتها أفضل في صورة المصفوفات عن الصياغة الجبرية التقليدية التي لم تعد كثيرة الاستخدام الآن.

9.C القيم المتوقعة المتنبأ بها : MEAN PREDICTION

دع:

$$\mathbf{X}_0 = \begin{bmatrix} 1 \\ X_{02} \\ X_{03} \\ \vdots \\ X_{0k} \end{bmatrix} \qquad (1.9.C)$$

تمثل متجهًا من قيم متغيرات الـ X والتي نرغب في استخدامه للتنبؤ بـ$\hat{Y}_0$ ، القيمة المتوسطة المتنبأ لـ Y .

الآن الانحدار المتعدد المقدر، في الشكل التقليدي، هو كالتالي :

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + u_i \qquad \textbf{(2.9.C)}$$

وفي صورة مصفوفات، فإن ذلك يمكن كتابته كالتالي :

$$\hat{Y}_i = \mathbf{x}_i' \hat{\boldsymbol{\beta}} \qquad \textbf{(3.9.C)}$$

حيث : $\mathbf{x}_i' = [1 \ X_{2i} \ X_{3i} \ \cdots \ X_{ki}]$ و

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

المعادلة : (2.9.C) أو (3.9.C) هي بالطبع القيمة المتوقعة المتنبأ بها لـ Y_i المرتبطة بقيمة $\mathbf{x}_i'$ المعلومة .

إذا كان $\mathbf{x}_i'$ كما هو معطى في (1.9.C)، فإن (3.9.C) تصبح كالتالي :

$$(\hat{Y}_i \mid \mathbf{x}_0') = \mathbf{x}_0' \hat{\boldsymbol{\beta}} \qquad \textbf{(4.9.C)}$$

حيث إن القيمة $\mathbf{x}_0$ بالطبع تعتبر قيمًا محددة. لاحظ أن (4.9.C) تعطي تنبؤاً غير متحيز لـ $E(Y_i \mid \mathbf{x}_0')$ حيث إن : $E(x_0'\hat{\boldsymbol{\beta}}) = \mathbf{x}_0'\hat{\boldsymbol{\beta}}$ (لماذا؟)

تباين القيمة المتوقعة المتنبأ بها : Variance of Mean Prediction

المعادلة الخاصة بتقدير تباين $(\hat{Y}_0 \mid \mathbf{x}_0')$ هو كالتالي[7] :

$$\operatorname{var}(\hat{Y}_0 \mid \mathbf{x}_0') = \sigma^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 \qquad \textbf{(5.9.C)}$$

حيث σ^2 هو تباين u_i، $\mathbf{x}_0'$ هو القيمة المعطاة للمتغيرات X التي نستخدمها في التنبؤ و$(\mathbf{X}'\mathbf{X})$ هو المصفوفة المعطاة في (9.3.C). في الواقع العملي، نستبدل σ^2 بالمقدر غير المتحيز $\hat{\sigma}^2$.

(7) لكيفية الاشتقاق، انظر J. Johnston, Econometrics Methods, McGraw-Hill, 3ded., New York, 1984, pp. 195–196.

سنستعرض أكثر القيم المتوقعة المتنبأ بها ، والتباين في الفقرة التالية :

التنبؤ بالمفردة : Individual prediction

قد أشرنا في الفصلين (5 و 8) أن التنبؤ بالمفردة لـ Y ($=Y_0$) معطى في (3.9.C) أو على وجه الخصوص (4.9.C). الفرق بين التنبؤ بالقيمة المتوقعة أو التنبؤ بمفردة يكمن في تبيانهما .

تباين التنبؤ بالمفردة : Variance of individual prediction

معادلة تباين تنبؤ المفردة هي كالتالي :

$$\operatorname{var}(Y_0 \mid \mathbf{x}_0) = \sigma^2[1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0] \qquad (6.9.C)$$

حيث $\operatorname{var}(Y_0 \mid \mathbf{x}_0)$ تعني $E(Y_0 - \hat{Y}_0 | X)^2$. عملياً تستبدل σ^2 بمقدرها غير المتحيز $\hat{\sigma}^2$. سنستعرض هذه المعادلة أكثر في الفقرة التالية :

10.C تلخيص أسلوب المصفوفات :
SUMMARY OF THE MATRIX APPROACH :

دعنا نستعرض مثالاً توضيحياً an illustrative example

البيانات المعطاة في جدول (4.C)، هذه البيانات بخصوص نفقات الاستهلاك الشخصي (PPCE)، و الدخل الشخصي المتغير (PPDI)، ومتغير الزمن أو الاتجاه العام. بإضافة متغير الاتجاه العام في النموذج، نحاول أن نفسر العلاقة بين PPCE و PPDI في إطار متغير الاتجاه العام (والذي قد يمثل العديد من المتغيرات الأخرى، مثل التطور التكنولوچي، التغير في الذوق، وهكذا).

ولتسهيل الجانب التطبيقي ، دعنا نفترض نموذج الانحدار التالي :

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{u}_i \qquad (1.10.C)$$

حيث Y = نفقات الاستهلاك الشخصي، X_2 = الدخل الشخصي المتغير و X_3 = الزمن. البيانات المطلوبة لهذا الانحدار (1.10.C) معطاة في جدول (4.C).

جدول (4.C) نفقات الاستهلاك الشخصي (PPCE) ، والدخل الشخصي المتغير (PPDI) في الولايات المتحدة ، 1956– 1970 بدولارات 1958

PPCE, Y	PPDI, X_2	Time, X_3	PPCE, Y	PPDI, X_2	Time, X_3
1673	1839	1 (= 1956)	1948	2126	9
1688	1844	2	2048	2239	10
1666	1831	3	2128	2336	11
1735	1881	4	2165	2404	12
1749	1883	5	2257	2487	13
1756	1910	6	2316	2535	14
1815	1969	7	2324	2595	15 (= 1970)
1867	2016	8			

المصدر: Economic Report of the president, January 1973 table B-16

باستخدام المصفوفات ، المشكلة السابقة يتم التعبير عنها كالتالي :

$$\begin{bmatrix} 1673 \\ 1688 \\ 1666 \\ 1735 \\ 1749 \\ 1756 \\ 1815 \\ 1867 \\ 1948 \\ 2048 \\ 2128 \\ 2165 \\ 2257 \\ 2316 \\ 2324 \end{bmatrix} = \begin{bmatrix} 1 & 1839 & 1 \\ 1 & 1844 & 2 \\ 1 & 1831 & 3 \\ 1 & 1881 & 4 \\ 1 & 1883 & 5 \\ 1 & 1910 & 6 \\ 1 & 1969 & 7 \\ 1 & 2016 & 8 \\ 1 & 2126 & 9 \\ 1 & 2239 & 10 \\ 1 & 2336 & 11 \\ 1 & 2404 & 12 \\ 1 & 2487 & 13 \\ 1 & 2535 & 14 \\ 1 & 2595 & 15 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} + \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \\ \hat{u}_5 \\ \hat{u}_6 \\ \hat{u}_7 \\ \hat{u}_8 \\ \hat{u}_9 \\ \hat{u}_{10} \\ \hat{u}_{11} \\ \hat{u}_{12} \\ \hat{u}_{13} \\ \hat{u}_{14} \\ \hat{u}_{15} \end{bmatrix} \qquad (2.10.C)$$

$$\underset{15\times 1}{\mathbf{y}} = \underset{15\times 3}{\mathbf{X}} \; \underset{3\times 1}{\hat{\boldsymbol{\beta}}} + \underset{15\times 1}{\hat{\mathbf{u}}}$$

من البيانات السابقة نحصل على المقادير التالية :

$$\bar{Y} = 1942.333 \qquad \bar{X}_2 = 2126.333 \qquad \bar{X}_3 = 8.0$$

$$\sum (Y_i - \bar{Y})^2 = 830,121.333$$

$$\sum (X_{2i} - \bar{X}_2)^2 = 1,103,111.333 \qquad \sum (X_{3i} - \bar{X}_3)^2 = 280.0$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2n} \\ X_{31} & X_{32} & X_{33} & \cdots & X_{3n} \end{bmatrix} \begin{bmatrix} 1 & X_{21} & X_{31} \\ 1 & X_{22} & X_{32} \\ 1 & X_{23} & X_{33} \\ \vdots & \vdots & \vdots \\ 1 & X_{2n} & X_{3n} \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} \\ \sum X_{3i} & \sum X_{2i}X_{3i} & \sum X_{3i}^2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & 31,895 & 120 \\ 31,895 & 68,922.513 & 272,144 \\ 120 & 272,144 & 1240 \end{bmatrix} \qquad \text{(3.10.C)}$$

$$\mathbf{X'y} = \begin{bmatrix} 29,135 \\ 62,905,821 \\ 247,934 \end{bmatrix} \qquad \text{(4.10.C)}$$

باستخدام قواعد معكوس المصفوفة الموجودة في الملحق B، نحصل على التالي:

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 37.232491 & -0.0225082 & 1.336707 \\ -0.0225082 & 0.0000137 & -0.0008319 \\ 1.336707 & -0.0008319 & 0.054034 \end{bmatrix} \qquad \text{(5.10.C)}$$

وبالتالي:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y} = \begin{bmatrix} 300.28625 \\ 0.74198 \\ 8.04356 \end{bmatrix} \qquad \text{(6.10.C)}$$

مجموع مربعات البواقي يمكن حسابه، كالتالي:

$$\begin{aligned} \sum \hat{u}_i^2 &= \hat{\mathbf{u}}'\hat{\mathbf{u}} \\ &= \mathbf{y'y} - \hat{\boldsymbol{\beta}}'\mathbf{X'y} \\ &= 57,420,003 - [300.28625 \quad 0.74198 \quad 8.04356] \begin{bmatrix} 29,135 \\ 62,905,821 \\ 247,934 \end{bmatrix} \\ &= 1976.85574 \end{aligned} \qquad \text{(7.10.C)}$$

وبالتالي نحصل على:

$$\hat{\sigma}^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{12} = 164.73797 \qquad \text{(8.10.C)}$$

مصفوفة التباين- التغاير لـ $\hat{\boldsymbol{\beta}}$ يمكن الحصول عليها كالتالي:

$$\text{var-cov}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2(\mathbf{X'X})^{-1} = \begin{bmatrix} 6133.650 & -3.70794 & 220.20634 \\ -3.70794 & 0.00226 & -0.13705 \\ 220.20634 & -0.13705 & 8.90155 \end{bmatrix} \qquad \text{(9.10.C)}$$

عناصر القطر الرئيسي لهذه المصفوفة يعطي تباين $\hat{\beta}_1$ ، $\hat{\beta}_2$ و $\hat{\beta}_3$ بالترتيب ، والجذر الموجب لهذه العناصر يعطي انحراف المعيار المناظر.

من البيانات السابقة، يمكن من السهل إثبات أن :

$$\text{ESS}: \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2 = 828{,}144.47786 \qquad \textbf{(10.10.C)}$$

$$\text{TSS}: \mathbf{y}'\mathbf{y} - n\bar{Y}^2 = 830{,}121.333 \qquad \textbf{(11.10.C)}$$

وبالتالي :

$$\begin{aligned} R^2 &= \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2} \\ &= \frac{828{,}144.47786}{830{,}121.333} \\ &= 0.99761 \end{aligned} \qquad \textbf{(12.10.C)}$$

بتطبيق (4.8.7) الخاصة بمعامل التحديد المعدل يمكن الحصول على :

$$\bar{R}^2 = 0.99722 \qquad \textbf{(13.10.C)}$$

وبتجميع كل النتائج السابقة، يكون لدينا :

$$\begin{aligned} \hat{Y}_i &= 300.28625 \;+\; 0.74198X_{2i} \;+\; 8.04356X_{3i} \\ &\quad (78.31763) \quad (0.04753) \quad (2.98354) \\ t &= \; (3.83421) \quad (15.60956) \quad (2.69598) \\ R^2 &= 0.99761 \qquad \bar{R}^2 = 0.99722 \qquad \text{df} = 12 \end{aligned} \qquad \textbf{(14.10.C)}$$

تفسير (14.10.C) كالتالي : إذا كان X_2 و X_3 يساويان الصفر ، فإن القيمة المتوسطة لنفقات الاستهلاك الشخصي مقدرة بـ 300 دولار. وكالمعتاد، هذا التفسير التقليدى للجزء الثابت المقطوع من المحور الصادي، يجب أن يتم تناوله بحذر شديد. معامل الانحدار الجزئي المساوي لـ 0.74198 يعني أنه، بافتراض ثبات باقي العوامل الأخرى، فإن كل زيادة في الدخل مثلاً، دولار واحد سيؤدي إلى زيادة في نفقات الاستهلاك الشخصي بمقدار 74 سنتًا تقريباً. باختصار، فإن الميل الحدي للاستهلاك مقدر بحوالي 0.74 أو 47 سنتًا تقريباً. بالمثل بافتراض ثبات باقي العوامل الأخرى، فإن متوسط نفقات الاستهلاك الشخصي يزيد بمعدل 8$ في السنة خلال فترة الدراسة، 1956–1970. قيم R^2 المساوية لـ 0.9976 توضح أن المتغيرين المفسرين الموجودين في النموذج يفسران حوالي 99 في المائة من تباين نفقات الاستهلاك في الولايات المتحدة خلال الفترة 1956–1970.

بالرجوع إلى المعنوية الإحصائية للمعاملات المقدرة ، نرى من (14.10.C) أن كلاً من المعاملات المقدرة له معنوية إحصائية فردياً عند، مثلاً، 5% مستوى معنوية: النسبة بين المعاملات المقدرة وأخطائها القياسية (أي، نسبة t) هي 3.8342، 15.6677 و 2.69598 بالترتيب. باستخدام اختبار t عند مستوى معنوية 5%، نجد أن قيمة t الحرجة بدرجة حرية 12 هي 2.179 . كل قيمة من قيم t المحسوبة تزيد عن القيمة الحرجة، وبالتالي، على المستوى الفردي، فإننا نرفض الفرض العدمي القائل بأن قيمة المجتمع الحقيقية للمعاملات السابقة تساوي الصفر.

كما لاحظنا من قبل ، لايمكننا استخدام اختبار t للفرض الخاص بأن $\beta_2 = \beta_3 = 0$ آنياً، حيث إن طريقة اختبار t تفترض أن العينة المسحوبة في كل مرة يطبق فيها اختبار t تكون عينة مستقلة. وإذا تم استخدام نفس العينة لاختبار الفرض الخاص بـ β_2 و β_3 آنياً فإنه من المحتمل أن يكون هناك ارتباط بين $\hat{\beta}_2$ و $\hat{\beta}_3$ مما يخالف فرض طريقة اختبار t الأساسي[9].

في واقع الأمر بالنظر إلى مصفوفة التباين – التغاير لـ$\hat{\beta}$ المعطاة في (9.10.C) نرى أن المقدرين $\hat{\beta}_2$ و $\hat{\beta}_3$ بينهما ارتباط سلبي (التغاير بينهما يساوي −0.13705). وبالتالي لانستطيع استخدام اختبار t لاختبار الفرض العدمي القائل بأن $\beta_2 = \beta_3 = 0$.

ولكن تذكر أن الفرض العدمي مثل $\beta_2 = \beta_3 = 0$، آنياً، يمكن اختباره بأسلوب تحليل التباين واختبار F، والذي سبق وشرحناه في الفصل (8). بالنسبة للمسألة الحالية، فإن جدول تحليل التباين معطى في جدول (5.C). وتحت الفروض التقليدية، نحصل على:

$$F = \frac{414,072.3893}{164.73797} = 2513.52 \qquad \textbf{(15.10.C)}$$

وهذه العينة تتبع توزيع F بدرجات حرية 2 و 12. قيمة F المحسوبة لها معنوية إحصائية عالية، وبالتالي نرفض الفرض العدمي القائل بأن $\beta_2 = \beta_3 = 0$ أي أن نفقات الاستهلاك الشخصي ليست مرتبطة خطياً مع الدخل الشخصي المتغير والاتجاه العام.

في الفقرة 9.C ناقشنا الطرق المختلفة للتنبؤ. سواء التنبؤ بالقيمة المتوقعة أو التنبؤ بالمفردة. افترض أنه في عام 1971 كان PPDI يساوي 2610$ وأردنا التنبؤ بـ PPCE المناظر لهذه القيمة. وبالتالي فإن القيمة المتوسطة المتنبأ بها والمفردة المتنبأ بها أيضاً لـ PPCE لعام 1971 لها قيمة واحدة وهي:

(9) انظر الفقرة 5.8 لمزيد من التفاصيل.

$$(\text{PPCE}_{1971} \mid \text{PPDI}_{1971}, X_3 = 16) = \mathbf{x}'_{1971}\hat{\boldsymbol{\beta}}$$
$$= [1 \quad 2610 \quad 16]\begin{bmatrix} 300.28625 \\ 0.74198 \\ 8.04356 \end{bmatrix} \qquad \textbf{(16.10.C)}$$
$$= 2365.55$$

حيث تم استخدام المعادلة (3.9.C)

تباين $\hat{Y}_{1971}$ و Y_{1971}، كما عرفنا من الفقرة 9.C، مختلفان وهما كالتالي:

$$\text{var}(\hat{Y}_{1971} \mid \mathbf{x}'_{1971}) = \hat{\sigma}^2[\mathbf{x}'_{1971}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{1971}]$$
$$= 164.73797[1 \quad 2610 \quad 16](\mathbf{X}'\mathbf{X})^{-1}\begin{bmatrix} 1 \\ 2610 \\ 16 \end{bmatrix} \qquad \textbf{(17.10.C)}$$

حيث $(\mathbf{X}'\mathbf{X})^{-1}$ هي معطاة في (5.10.C). بالتعويض عنها في (17.10.C) يمكن للقارئ أن يثبت أن:

$$\text{var}(\hat{Y}_{1971} \mid \mathbf{x}'_{1971}) = 48.6426 \qquad \textbf{(18.10.C)}$$

جدول (5.C) جدول **ANOVA** لبيانات جدول (4.C)

Source of variation	SS	df	MSS
Due to X_2, X_3	828,144.47786	2	414,072.3893
Due to residuals	1,976.85574	12	164.73797
Total	830,121.33360	14	

وبالتالي :

$$\text{se}(\hat{Y}_{1971} \mid \mathbf{x}'_{1971}) = 6.9744$$

وسنترك للقارئ إثبات، باستخدام (6.9.C)، أن

$$\text{var}(Y_{1971} \mid \mathbf{x}'_{1971}) = 213.3806 \qquad \textbf{(19.10.C)}$$

و

$$\text{se}(Y_{1971} \mid \mathbf{x}'_{1971}) = 14.6076$$

لاحظ أن: $\text{var}(Y_{1971} \mid \mathbf{x}'_{1971}) = E[Y_{1971} - \hat{Y}_{1971} \mid \mathbf{x}'_{1971}]^2$

في الفقرة 5.C تناولنا مصفوفة الارتباط R. بالنسبة للبيانات الخاصة بالمسألة الحالية فإن مصفوفة الارتباط كالتالي:

$$
R = \begin{array}{c} \\ Y \\ X_2 \\ X_3 \end{array}
\begin{array}{c}
\begin{array}{ccc} Y & X_2 & X_3 \end{array} \\
\begin{bmatrix} 1 & 0.9980 & 0.9743 \\ 0.9980 & 1 & 0.9664 \\ 0.9743 & 0.9664 & 1 \end{bmatrix}
\end{array}
\qquad (20.10.C)
$$

لاحظ أنه في (20.10.C) قد ألحقنا بمصفوفة الارتباط المتغيرات الموجودة في النموذج، حيث تستطيع بسهولة تحديد المتغيرات المتعلقة بمعاملات الارتباط، وبالتالي فالمعامل 0.998 الموجود في الصف الأول للمصفوفة (20.10.C) معامل الارتباط بين Y و X_2 (أي r_{12}). من ارتباط الدرجة صفر المعطى في مصفوفة الارتباط (20.10.C) يمكن بسهولة اشتقاق معاملات الارتباط من الدرجة الأولى (انظر تمرين 7.C).

11.C المربعات الصغرى العامة :
GENERALIZED LEAST SQUARES (GLS)

في العديد من المرات ذكرنا أن OLS هي حالة خاصة من GLS. لرؤية ذلك دعنا نعود إلى المعادلة (2.2.C). ولكن نضع في الاعتبار عدم ثبات التباين [العناصر الموجودة على القطر الرئيسي لـ (2.2.C)] والارتباط الذاتي الموجود في مقادير الأخطاء [العناصر الموجودة على جانبي القطر الرئيسي لـ (2.2.C)] دعنا نفترض أن:

$$\mathrm{E}(\mathbf{uu}') = \sigma^2\mathbf{V} \qquad (1.11.C)$$

حيث $\mathbf{V}$ هي مصفوفة $n \times n$ معلومة.

وبالتالي إذا كان النموذج هو:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$$

حيث $E(\mathbf{u}) = \mathbf{0}$ و $\text{var-cov}(\mathbf{u}) = \sigma^2\mathbf{V}$. في حالة σ^2 غير معلومة، وهذه هي الحالة الموجودة الآن، $\mathbf{V}$ تمثل التباين والتغاير المفترض بين الخطأ العشوائي u_t.

باستخدام فروض التباين - التغاير الخاصة بمقادير الأخطاء، يمكن إثبات أن:

$$\beta^{gls} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \qquad (2.11.C)$$

β^{gls} معروفة باسم مقدر المربعات الصغرى العام (GLS) لـ β، ويمكن إثبات أيضاً أن:

$$\text{var-cov}(\beta^{gls}) = \sigma^2\,(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \qquad (3.11.C)$$

ويمكن إثبات أن β^{gls} هو أفضل مقدر خطي غير متحيز لـ β.

إذا فرضنا أن تباين كل مقدار من مقادير الأخطاء يساوي ثابت σ^2، وكانت مقادير الأخطاء غير مرتبطة تبادلياً، فإن المصفوفة **V** تصبح مصفوفة الوحدة، كما هو موضح في (3.2.C). إذا كانت مقادير الأخطاء غير مرتبطة تبادلياً، ولكن مختلفة في التباين (أي توجد ظاهرة اختلاف التباين)، فإن المصفوفة **V** ستكون مصفوفة قطرية، عناصر قطرها الرئيسي تمثل التباينات المختلفة. وبالطبع إذا كان هناك اختلاف في التباين وارتباط ذاتي، فإن المصفوفة **V** سيكون لديها عناصر على القطر الرئيسي، وعناصر أيضاً على جانب القطر الرئيسي.

المشكلة الأساسية في الواقع، هي أننا لانعرف قيمة σ^2 ولاقيمة التباينات ولا التغايرات (أي المصفوفة **V**). وكحل لذلك، فإننا نستطيع استخدام طريقة المربعات الصغرى العامة المقدرة. أولاً سنقدر نموذجنا باستخدام OLS بغض النظر عن مشكلة اختلاف التباين أو الارتباط الذاتي. قد حصلنا على بواقي هذا النموذج ومصفوفة التباين- التغاير (المقدرة) لمقدار الخطأ بالتعويض عن كل العناصر المجهولة السابقة في (2.2.C) بالقيمة المقدرة لـ u، أي $\hat{u}$. يمكن إثبات أن مقدرات EGLS هي مقدرات متسقة لـ GLS ورمزياً ذلك يعني أن:

$$\beta^{egls} = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{y}) \qquad (4.11.C)$$

$$\text{var-cov}(\beta^{egls}) = \sigma^2(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1} \qquad (5.11.C)$$

حيث: $\hat{\mathbf{V}}$ هي مقدرات لـ **V**.

12.C الخلاصة والنتائج : SUMMARY AND CONCLUSIONS

الهدف الرئيسي لهذا الملحق، هو تقديم نموذج الانحدار الخطي من خلال المصفوفات. فعلى الرغم من استخدام القليل من المفاهيم الجديدة في تحليل الانحدار، إلا أن أسلوب المصفوفات يعتبر صورة مختصرة للتعامل مع نماذج الانحدار الخطي التي تحتوي على أي عدد من المتغيرات.

في نهاية هذا الملحق، لاحظ أنه إذا كانت المتغيرات X و Y مقاسة في شكل انحرافات، أي انحرافات عن الوسط الحسابي، فهناك بعض التغيرات المحدودة في المعادلات التي تم تقديمها من قبل. هذه التغيرات مقدمة في جدول (6.C).

جدول (6.C) نموذج انحدار لـ k متغير بوحداتهم الأصلية وفي شكل انحرافات(*)

Original units		Deviation form	
$\mathbf{y} = \mathbf{X}\hat{\boldsymbol\beta} + \hat{\mathbf{u}}$	(C.3.2)	$\mathbf{y} = \mathbf{X}\hat{\boldsymbol\beta} + \hat{\mathbf{u}}$ The column of 1's in the **X** matrix drops out. (Why?)	
$\hat{\boldsymbol\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$	(C.3.11)	Same	
$\text{var-cov}(\hat{\boldsymbol\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$	(C.3.13)	Same	
$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol\beta}'\mathbf{X}'\mathbf{y}$	(C.3.18)	Same	
$\sum y_i^2 = \mathbf{y}'\mathbf{y} - n\bar{Y}^2$	(C.3.16)	$\sum y_i^2 = \mathbf{y}'\mathbf{y}$	(C.12.1)
$\text{ESS} = \hat{\boldsymbol\beta}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2$	(C.3.17)	$\text{ESS} = \boldsymbol\beta'\mathbf{X}'\mathbf{y}$	(C.12.2)
$R^2 = \dfrac{\hat{\boldsymbol\beta}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}$	(C.4.2)	$R^2 = \dfrac{\hat{\boldsymbol\beta}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}$	(C.12.3)

(*) لاحظ أنه على الرغم من أن رموز المصفوفات والمتجهات واحدة في الحالتين، إلا أنه في حالة الانحرافات فإن عناصر المصفوفات والمتجهات يفترض انها انحرافات بدلاً من قيم أصلية. لاحظ أيضاً أن انحرافات $\hat{\beta}$ من الدرجة $k-1$ و $\text{var-cov}(\hat{\beta})$ من الدرجة $(k-1)(k-1)$

من هذا الجدول[10] يتضح أنه باستخدام الانحرافات، فإن تصحيح المتوسط $n\bar{Y}^2$ يحذف من الـ TSS و ESS. (لماذا؟) هذا الانخفاض يؤدي إلى تغير في معادلة R^2. وبخلاف ذلك، فإن معظم المعادلات تظل في قياساتها الأصلية حتى باستخدام الانحرافات.

تمارين: EXERCISES

1.C للمثال التوضيحي الموجود في الفقرة 10.C فإن $\mathbf{X}'\mathbf{X}$ و $\mathbf{X}'\mathbf{y}$ باستخدام البيانات في صورة انحرافات تكون كالتالي:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1{,}103{,}111.333 & 16{,}984 \\ 16{,}984 & 280 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 955{,}099.333 \\ 14{,}854.000 \end{bmatrix}$$

(**a**) قدر β_2 و β_3.

(**b**) كيف تقدر β_1؟

(**c**) احصل على تباين $\hat{\beta}_2$ و $\hat{\beta}_3$ وتباينهما.

(10) في عصر التطور الكبير الحادث في الحاسبات الآلية قد لاتحتاج إلى شكل الانحرافات. ولكن ذلك يبسط المعادلات وبالتالي الحسابات.

(d) احصل على R^2 و $\bar{R}^2$.

(e) بمقارنة نتائجك مع نظيرها الذي حصلنا عليه في الفقرة 10.C . ماذا تجد من مميزات لاستخدام الانحرافات؟

2.C بالعودة إلى تمرين 22.23 . باستخدام البيانات المعطاة كون المصفوفة (X′X) المناسبة والمتجه X′y وقدر متجه المعلمات β ومصفوفة التباين والتغاير الخاصة به. احصل أيضاً على R^2 كيف يمكنك اختبار فرض مرونة M 1 بالنسبة لـ GDP وبالنسبة لمعدل الفائدة R متساوية؟

3.C اختبار تساوي معاملين انحداريين . افترض أن لديك نموذج الانحدار التالي :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

وتريد اختبار الفرض القائل بأن $\beta_2 = \beta_3$. إذا افترضنا أن u_i لها التوزيع الطبيعي ، فيمكن إثبات أن :

$$t = \frac{\hat{\beta}_2 - \hat{\beta}_3}{\sqrt{\text{var}(\hat{\beta}_2) + \text{var}(\hat{\beta}_3) - 2\,\text{cov}(\hat{\beta}_2, \hat{\beta}_3)}}$$

يتبع توزيع t بدرجات حرية تساوي $n - 3$ (انظر الفقرة 8.6) . (عموماً، في حالة وجود k متغير، فإن df تكون $n - k$) وبالتالي اختبار t السابق يمكن استخدامه لاختبار الفرض العدمي $\beta_2 = \beta_3$.

طبق اختبار t السابق لاختبار الفرض القائل بأن القيم الحقيقية لـ β_2 و β_3 في انحدار (14.10.C) متساويان .

ملحوظة مساعدة : استخدم مصفوفة التباين – التغاير لـ β المعطاة في (9.10.C) .

4.C التعبير عن الارتباط بدرجة عليا في صورة ارتباط بدرجة أقل . معامل الارتباط من الدرجة p يمكن التعبير عنه في صورة معامل ارتباط بدرجة $p - 1$ باستخدام صيغة الاختزال التالية :

$$r_{12.345\ldots p} = \frac{r_{12.345\ldots(p-1)} - [r_{1p.345\ldots(p-1)} r_{2p.345\ldots(p-1)}]}{\sqrt{[1 - r^2_{1p.345\ldots(p-1)}]}\sqrt{[1 - r^2_{2p.345\ldots(p-1)}]}}$$

وبالتالي :

$$r_{12.3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{1 - r^2_{13}}\sqrt{1 - r^2_{23}}}$$

كما في الفصل (7).

افترض أن لديك مصفوفة الارتباط التالية:

$$\mathbf{R} = \begin{matrix} & Y & X_2 & X_3 & X_4 & X_5 \\ Y & 1 & 0.44 & -0.34 & -0.31 & -0.14 \\ X_2 & & 1 & 0.25 & -0.19 & -0.35 \\ X_3 & & & 1 & 0.44 & 0.33 \\ X_4 & & & & 1 & 0.85 \\ X_5 & & & & & 1 \end{matrix}$$

أوجد التالي:

a. $r_{12.345}$ b. $r_{12.34}$ c. $r_{12.3}$

d. $r_{13.245}$ e. $r_{13.24}$ f. $r_{13.2}$

5.C التعبير عن معاملات الانحدار من درجة عليا في صورة معاملات انحدار أقل.
معامل الانحدار من الدرجة p يمكن التعبير عنه في صورة معامل انحدار من الدرجة $p-1$ باستخدام صيغة الاختزال التالية:

$$\hat{\beta}_{12.345\ldots p} = \frac{\hat{\beta}_{12.345\ldots(p-1)} - [\hat{\beta}_{1p.345\ldots(p-1)}\hat{\beta}_{p2.345\ldots(p-1)}]}{1 - \hat{\beta}_{2p.345\ldots(p-1)}\hat{\beta}_{p2.345\ldots(p-1)}}$$

وبالتالي:

$$\hat{\beta}_{12.3} = \frac{\hat{\beta}_{12} - \hat{\beta}_{13}\hat{\beta}_{32}}{1 - \hat{\beta}_{23}\hat{\beta}_{32}}$$

حيث $\beta_{12.3}$ هو معامل الميل في انحدار y على X_2، وبافتراض ثبات X_3. وبالمثل $\beta_{12.94}$ هو معامل ميل انحدار Y على X_2 بافتراض ثبات X_3 و X_4 وهكذا.
باستخدام المعادلة السابقة، أوجد صيغة للتعبير عن معاملات الانحدار التالية في صورة معاملات انحدار من درجة أقل: $\hat{\beta}_{12.3456}$ ، $\hat{\beta}_{12.345}$ و $\hat{\beta}_{12.34}$.

6.C تحقق من المعادلة التالية:

$$\hat{\beta}_{12.3}\hat{\beta}_{23.1}\hat{\beta}_{31.2} = r_{12.3}r_{23.1}r_{31.2}$$

7.C أوجد كل معاملات الارتباط الجزئية من الدرجة الأولى لمصفوفة الارتباط R المعطاة في (20.10.C).

8.C في دراسة لاختلاف معدلات الجريمة في مدن ما كبيرة في الولايات المتحدة ogburn حصل على البيانات التالية(*):

(*) W.F. Ogburn, "Factors in the Variation of Crime among Cities, "Journal of American Statistical Association, vol. 30, 1935, p. 12.

$$\bar{Y} = 19.9 \quad S_1 = 7.9$$
$$\bar{X}_2 = 49.2 \quad S_2 = 1.3$$
$$\bar{X}_3 = 10.2 \quad S_3 = 4.6$$
$$\bar{X}_4 = 481.4 \quad S_4 = 74.4$$
$$\bar{X}_5 = 41.6 \quad S_5 = 10.8$$

$$\mathbf{R} = \begin{array}{c} \\ Y \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{array} \begin{array}{c} \begin{array}{ccccc} Y & X_2 & X_3 & X_4 & X_5 \end{array} \\ \begin{bmatrix} 1 & 0.44 & -0.34 & -0.31 & -0.14 \\ & 1 & 0.25 & -0.19 & -0.35 \\ & & 1 & 0.44 & 0.33 \\ & & & 1 & 0.85 \\ & & & & 1 \end{bmatrix} \end{array}$$

حيث : Y = معدل الجريمة، عدد التهم المعروفة لكل ألف نسمة من السكان

X_2= نسبة الذكور الذين يقومون بالجريمة

X_3= نسبة الذكور الأجانب الذين يقومون بالجريمة

X_4= عدد الأطفال أقل من 5 سنوات لكل ألف سيدة متزوجة عمرها بين 14 و44 سنة.

X_5= العضوية في كنيسة ما، العضوية في الكنيسة لمدة 13 عامًا أو أكثر لكل 100 من السكان أعمارهم 13 سنة أو أكثر، S_1 إلى S_5 هي الانحرافات المعيارية من العينة للمتغيرات Y إلى X_5 و **R** هي مصفوفة الارتباط.

(**a**) باعتبار Y متغير تابع، احصل على انحدار Y على الأربعة متغيرات X وفسر الانحدار المقدر.

(**b**) احصل على $r_{12.3}$، $r_{14.35}$ و $r_{5.34}$.

(**c**) احصل على R^2 واختبر الفرض القائل بأن كل معاملات الميل الجزئية آنياً تساوي الصفر.

9.C الجدول التالي يحتوي على بيانات الناتج وتكلفة إنتاج سلعة ما في المدى القصير (انظر مثال 4.7).

Output	Total cost, $
1	193
2	226
3	240
4	244
5	257
6	260
7	274
8	297
9	350
10	420

لاختبار ما إذا كانت البيانات السابقة تقترح متوسط له الشكل U ومنحنيات تكلفة حدية متطابقة في المدى القصير يمكن استخدام النموذج التالي :

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \beta_4 X_i^3 + u_i$$

حيث Y = التكلفة الكلية، وX = ناتج المتغيرات المفسرة الإضافية X_i^2 وX_i^3 مأخوذة من X.

(a) عبر عن البيانات في صورة انحرافات واحصل على $(\mathbf{X'X})$، $(\mathbf{X'y})$ و $(\mathbf{X'X})^{-1}$.

(b) قدر β_2، β_3 و β_4.

(c) قدر مصفوفة var-cov لـ $\hat{\boldsymbol{\beta}}$

(d) قدر β_1. فسر $\hat{\beta}_1$ في إطار المسألة الحالية.

(e) احصل على R^2 و$\bar{R}^2$.

(f) مبدئياً ماهي إشارة β_2، β_3 و β_4؟ لماذا؟

(g) من دالة التكلفة الكلية المعطاة سابقاً، احصل على صيغة دالة التكلفة الحدية، ودالة التكلفة المتوسطة.

(h) وفق دالة التكلفة الحدية والمتوسطة للبيانات وعلق على هذا التوفيق.

(i) إذا كان $\beta_3 = \beta_4 = 0$، ماهي طبيعية دالة التكلفة الحدية؟ كيف تختبر الفرض القائل بأن $\beta_3 = \beta_4 = 0$؟

(j) كيف يمكنك اشتقاق دوال التكلفة الكلية المتغيرة، ودوال التكلفة المتوسطة المتغيرة من البيانات المعطاة؟

10.C لدراسة المشاركة في قوة العمل من العائلات غير الريفية (عائلات تكسب أقل من 3943$ في 1969)، تم الحصول على البيانات الموجودة في جدول (7.C) تم الحصول عليها من تعداد السكان لعام 1970 .

(a) باستخدام نموذج الانحدار $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i$، احصل على مقدرات معاملات الانحدار وفسر نتائجك.

(b) مبدئياً . . ماذا تتوقع بالنسبة لإشارات معاملات الانحدار في النموذج السابق ولماذا؟

جدول (7.C) حالة المشاركة في قوة العمل من العائلات الريفية الفقيرة : تعداد السكان ، مدينة New York ، 1970

Tract no.	% in labor force, Y^*	Mean family income, $X_2^\dagger$	Mean family size, X_3	Unemployment rate, $X_4^\ddagger$
137	64.3	1,998	2.95	4.4
139	45.4	1,114	3.40	3.4
141	26.6	1,942	3.72	1.1
142	87.5	1,998	4.43	3.1
143	71.3	2,026	3.82	7.7
145	82.4	1,853	3.90	5.0
147	26.3	1,666	3.32	6.2
149	61.6	1,434	3.80	5.4
151	52.9	1,513	3.49	12.2
153	64.7	2,008	3.85	4.8
155	64.9	1,704	4.69	2.9
157	70.5	1,525	3.89	4.8
159	87.2	1,842	3.53	3.9
161	81.2	1,735	4.96	7.2
163	67.9	1,639	3.68	3.6

Y^* = رب الأسرة أقل من 65 عامًا

$^\dagger X_2$ = الدولار

$^\ddagger X_4$ = نسبة قوة العمل المدنية غير العاملة

المصدر : Census Tracts: New York, Bureau of the Census, U.S. Department of Commerce. 1970.

(c) كيف يمكنك اختبار الفرض القائل بأن معدل البطالة الكلي ليس له تأثير على المشاركة في قوة العمل للعائلات الريفية الفقيرة في التعداد المعطى في الجدول السابق.

(d) هل يمكن إسقاط أو حذف أي متغير من النموذج السابق؟ لماذا؟

(e) ماهي المتغيرات الأخرى التي يمكن أن تدخلها في هذا النموذج؟

11.C في تطبيق لدالة إنتاج Cobb- Douglas النتائج التالية تم الحصول عليها :

$$\widehat{\ln Y_i} = 2.3542 + 0.9576 \ln X_{2i} + 0.8242 \ln X_{3i}$$
$$(0.3022) \qquad (0.3571)$$
$$R^2 = 0.8432 \qquad df = 12$$

حيث Y = الناتج ، و X_2= العمالة ، X_3= رأس المال ، والأرقام بين الأقواس هي الأخطاء القياسية المقدرة .

(a) كما لاحظنا من قبل في الفصل (7)، فإن معاملات قوة العمل ، ورأس المال في المعادلة السابقة تمثل مرونة الناتج بالنسبة للعمالة ورأس المال . اختبر الفرض القائل بأن هذه المرونات مساوية تماماً للوحدة الواحدة .

(**b**) اختبر الفرض القائل بأن مرونة العمالة تساوي مرونة رأس المال بافتراض (*i*) التغاير بين العمالة المقدرة ورأس المال المقدر تساوي الصفر و (*ii*) تساوي -0.0972 .

(**c**) كيف يمكن اختيار المعنوية الكلية لمعادلة الانحدار السابقة؟

12.C(*) عبر عن دالة الإمكان لنموذج الانحدار الخاص بالـ k متغير باستخدام المصفوفات، واثبت أن $\hat{\beta}$، متجه مقدرات الإمكان الأعظم، مساوي تماماً لـ $\hat{\beta}$، متجه مقدرات الـ OLS لنموذج انحدار الـ k متغير.

13.C الانحدار باستخدام متغيرات قياسية. اعتبر دوال انحدار العينة التالية (SRFs):

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{u}_i \quad (1)$$

$$Y_i^* = b_1 + b_2 X_{2i}^* + b_3 X_{3i}^* + \hat{u}_i^* \quad (2)$$

حيث :

$$Y_i^* = \frac{Y_i - \bar{Y}}{s_Y}$$

$$X_{2i}^* = \frac{X_{2i} - \bar{X}_2}{s_2}$$

$$X_{3i}^* = \frac{X_{3i} - \bar{X}_3}{s_3}$$

حيث الـ s's ترمز إلى الانحرافات المعيارية للعينة - كما سبق وأشرنا في الفصل 6 - الفقرة 3.6، فإن المتغيرات المعطاة أعلى معرفة بالمتغيرات القياسية. هذه المتغيرات لها متوسط يساوي الصفر، وانحراف معياري يساوي الوحدة (1=).

بوضع كل المتغيرات في صورة انحرافات، اثبت التالي بالنسبة للنموذج (2):

a. $\mathbf{X'X} = \begin{bmatrix} 1 & r_{23} \\ r_{23} & 1 \end{bmatrix} n$

b. $\mathbf{X'y} = \begin{bmatrix} r_{12} \\ r_{13} \end{bmatrix} n$

c. $\mathbf{X'X}^{-1} = \dfrac{1}{n(1 - r_{23}^2)} \begin{bmatrix} 1 & -r_{23} \\ -r_{23} & 1 \end{bmatrix}$

d. $\hat{\boldsymbol{\beta}} = \begin{bmatrix} b_2 \\ b_3 \end{bmatrix} = \dfrac{1}{1 - r_{23}^2} \begin{bmatrix} r_{12} - r_{23} r_{13} \\ r_{13} - r_{23} r_{12} \end{bmatrix}$

e. $b_1 = 0$

(*) اختياري.

اوجد العلاقة بين الـ b's والـ $\hat{\beta}$'s.

(لاحظ أنه في العلاقات السابقة، فإن n ترمز إلى حجم العينة، r_{12}، r_{13} و r_{23} ترمز إلى معاملات الارتباط بين Y و X_2، وبين Y و X_3، وبين X_2 و X_3 بالترتيب).

14.C اثبت العلاقات بين (15.10.C) و (19.10.C)

15.C (*) المربعات الصغرى المقيدة. افترض أن:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \qquad (1)$$

ونحن نريد التقدير وفقاً لمجموعة من القيود المتساوية التالية:

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r} \qquad (2)$$

حيث **R**= مصفوفة معلومة من الدرجة $q \times k$ $(q \leq k)$، و **r** هو متجه معلوم من q عنصر. للتوضيح دعنا نفترض أن نموذجنا هو:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_i \qquad (3)$$

وافترض أننا نريد تقدير هذا النموذج وفقاً لهذه القيود التالية:

$$\begin{aligned} \beta_2 - \beta_3 &= 0 \\ \beta_4 + \beta_5 &= 1 \end{aligned} \qquad (4)$$

يمكن أن نستخدم بعض الأساليب التي ناقشناها من قبل في الفصل (8) للتعامل مع هذه القيود (مثل $\beta_2 = \beta_3$ و $\beta_4 = 1 - \beta_5$، وبالتالي حذف β_2 و β_4 من النموذج) ونختبر مدى صحة هذه القيود باستخدام اختبار F الذي ناقشناه من قبل. ولكن كطريقة أكثر مباشرة في التقدير (3) ادخل القيود (4) مباشرة في عملية التقدير عن طريق التعبير أولاً عن القيود في شكل المعادلة (2) والذي يصبح كالتالي:

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad \mathbf{r} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad (5)$$

باعتبار $\beta^{(*)}$ تمثل مقدر المربعات الصغرى المقيد، يمكن أن نثبت أن β^* يمكن تقديرها من المعادلة التالية(*):

$$\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \qquad (6)$$

(*) انظر في J. Johnston, op. cit., p. 205.

حيث $\hat{\beta}$ هو المقدر العادي (غير الـمقيد) من المعادلة العادية $(X'X)^{-1} X'y$.

(a) ماهو المتجه β في (3)؟

(b) بمعلومية المتجه β، اثبت أن المصفوفة **R** والمتجه **r** المعطى في (5) متماشى مع القيود الموجودة في (4).

(c) أوجد **R** و **r** في الحالات التالية:

(i) $\beta_2 = \beta_3 = \beta_4 = 2$

(ii) $\beta_2 = \beta_3$ و $\beta_4 = \beta_5$

(iii) $\beta_2 - 3\beta_3 = 5\beta_4$

(iv) $\beta_2 + 3\beta_3 = 0$

(d) متى سيتحقق التالي $\hat{\beta}^* = \hat{\beta}$ ؟

ملحق AC — APPENDIX CA

1.AC اشتقاق المعادلات الطبيعية أو الآنية k :
DERIVATION OF k NORMAL OR SIMULTANEOUS EQUATIONS

بتفاضل :

$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})^2$$

جزيئياً بالنسبة لـ $\hat{\beta}_1$، $\hat{\beta}_2$ ، $\hat{\beta}_k$ نحصل على:

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_1} = 2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})(-1)$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_2} = 2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_k X_{ki})(-X_{2i})$$

....

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_k} = 2 \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{ki} - \cdots - \hat{\beta}_k X_{ki})(-X_{ki})$$

وبمساواة التفاضلات الجزئية السابقة إلى الصفر، وإعادة ترتيب مقاديرها نحصل على المعادلات الطبيعية k المعطاة في (8.3.C) .

2.AC اشتقاق المعادلات الطبيعية باستخدام المصفوفات :
MATRIX DERIVATION OF NORMAL EQUATIONS

من (7.3.C) نحصل على :

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

باستخدام قواعد تفاضل المصفوفات المعطاة في الملحق B، نحصل على :

$$\frac{\partial(\hat{\mathbf{u}}'\hat{\mathbf{u}})}{\partial\hat{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

بمساواة المعادلة السابقة مع الصفر نحصل على :

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

وبالتالي : $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{y}$ بشرط إمكانية وجود معكوس المصفوفة .

3.AC مصفوفة التباين - التغاير لـ $\hat{\beta}$:
VARIANCE- COVARIANCE MATRIX OF $\hat{\beta}$

من (11.3.C) نحصل على :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

بالتعويض عن : $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ في المعادلة السابقة نحصل على :

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{u})\\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\\ &= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\end{aligned} \tag{1}$$

وبالتالي :

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \tag{2}$$

بالتعريف :

$$\begin{aligned}\text{var-cov}\,(\hat{\boldsymbol{\beta}}) &= E[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})']\\ &= E\{[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}]'\}\\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]\end{aligned} \tag{3}$$

حيث في الخطوة السابقة تم استخدام $(\mathbf{AB'}) = \mathbf{B'A'}$

لاحظ أن : X's غير عشوائية ، وبإدخال التوقع على (3) نحصل على :

$$\begin{aligned}\text{var-cov}(\hat{\beta}) &= (\mathbf{X'X})^{-1}\mathbf{X'}E(\mathbf{uu'})\mathbf{X}(\mathbf{X'X})^{-1}\\ &= (\mathbf{X'X})^{-1}\mathbf{X'}\sigma^2\mathbf{IX}(\mathbf{X'X})^{-1}\\ &= \sigma^2(\mathbf{X'X})^{-1}\end{aligned}$$

وهذه هي النتيجة المعطاة في (13.3.C). لاحظ أنه عند اشتقاق النتيجة السابقة استخدمنا فرض $E(\mathbf{uu'})= \sigma^2\mathbf{I}$.

4.AC خاصية BLUE لمقدرات OLS :
BLUE PROPERTY OF OLS ESTIMATORS

من (11.3.C) لدينا :

$$\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'y} \qquad (1)$$

وبما أن : $(\mathbf{X'X})^{-1}\,\mathbf{X'}$ هو مصفوفة من أعداد ثابتة ، $\hat{\beta}$ هو دالة خطية في Y ، وبالتالي بالتعريف هو أيضاً مقدر خطي .

تذكر أن الـPRF هو :

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u} \qquad (2)$$

بالتعويض عن ذلك في (1)، نحصل على

$$\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'}(\mathbf{X}\beta + \mathbf{u}) \qquad (3)$$

$$= \beta + (\mathbf{X'X})^{-1}\mathbf{X'u} \qquad (4)$$

بما أن : $(\mathbf{X'X})^{-1}\,\mathbf{X'X} = \mathbf{I}$

بإدخال التوقع على (4) نحصل على :

$$\begin{aligned}E(\hat{\beta}) &= E(\beta) + (\mathbf{X'X})^{-1}\mathbf{X'}E(\mathbf{u})\\ &= \beta\end{aligned} \qquad (5)$$

بما أن : $E(\beta) = \beta$ (لماذا؟) و $E(\mathbf{u})= 0$ وفقاً للفروض الخاصة بالنموذج، مما يعني أن $\hat{\beta}$ مقدر غير متحيز لـ β .

دع $\hat{\beta}^*$ مقدر خطي آخر لـ β، يمكن كتابته كالتالي:

$$\beta^* = [(\mathbf{X'X})^{-1}\mathbf{X'} + \mathbf{C}]\mathbf{y} \qquad (6)$$

حيث C هي مصفوفة ثوابت

بالتعويض عن y من (2) في (6) نحصل على:

$$\begin{aligned}\hat{\beta}^* &= [(\mathbf{X'X})^{-1}\mathbf{X'} + \mathbf{C}](\mathbf{X}\beta + \mathbf{u}) \\ &= \beta + \mathbf{CX}\beta + (\mathbf{X'X})^{-1}\mathbf{X'u} + \mathbf{Cu}\end{aligned} \qquad (7)$$

الآن إذا كان $\hat{\beta}^*$ مقدر غير متحيز لـ β، لابد أن يكون لدينا

$$\mathbf{CX} = 0 \qquad \text{(لماذا؟)} \qquad (8)$$

باستخدام (7)، (8) يمكن كتابته كالتالي:

$$\hat{\beta}^* - \beta = (\mathbf{X'X})^{-1}\mathbf{X'u} + \mathbf{Cu} \qquad (9)$$

بالتعريف فإن var-cov ($\hat{\beta}^*$) هو

$$E(\hat{\beta}^* - \beta)(\hat{\beta}^* - \beta)' = E[(\mathbf{X'X})^{-1}\mathbf{X'u} + \mathbf{Cu}][(\mathbf{X'X})^{-1}\mathbf{X'u} + \mathbf{Cu}]' \quad (10)$$

باستخدام خاصية معكوس المصفوفة ومدورها وبعد بعض التبسيطات الجبرية، نحصل على

$$\begin{aligned}\text{var-cov}\,(\hat{\beta}^*) &= \sigma^2(\mathbf{X'X})^{-1} + \sigma^2\mathbf{CC'} \\ &= \text{var-cov}\,(\hat{\beta}) + \sigma^2\mathbf{CC'}\end{aligned} \qquad (11)$$

والذي يوضح أن مصفوفة التباين - التغاير للمقدر البديل غير المتحيز $\hat{\beta}^*$ تساوي مصفوفة التباين - التغاير لمقدر OLS $\hat{\beta}$ بالإضافة إلى σ^2 مضروبة في $\mathbf{CC'}$ والتي تعتبر مصفوفة موجبة شبه معرفة(*). وبالتالي تباين أي عنصر في $\hat{\beta}^*$ لابد أن يتساوى مع أو يكون أكبر من نظيره في $\hat{\beta}$ ، مما يعني أن $\hat{\beta}$ هو BLUE. بالطبع إذا كانت C هي المصفوفة الصفرية، أي أن C=0، فإن $\hat{\beta} = \hat{\beta}^*$ ، مما يعني أنه إذا وجدنا مقدر BLUE فإنه يجب أن يكون مقدر المربعات الصغرى $\hat{\beta}$.

(*) انظر المراجع في ملحق B.

ملحق

APPENDIX D — D

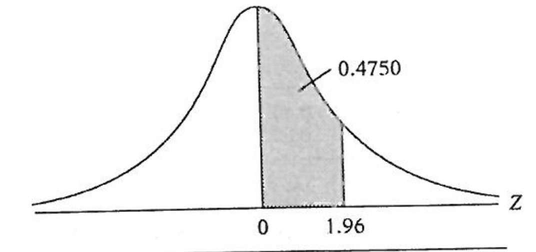
STATISTICAL TABLES — جداول إحصائية

TABLE D.1 AREAS UNDER THE STANDARDIZED NORMAL DISTRIBUTION

Example

$\Pr(0 \le Z \le 1.96) = 0.4750$

$\Pr(Z \ge 1.96) = 0.5 - 0.4750 = 0.025$



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4454	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Note: This table gives the area in the right-hand tail of the distribution (i.e., $Z \ge 0$). But since the normal distribution is symmetrical about $Z = 0$, the area in the left-hand tail is the same as the area in the corresponding right-hand tail. For example, $P(-1.96 \le Z \le 0) = 0.4750$. Therefore, $P(-1.96 \le Z \le 1.96) = 2(0.4750) = 0.95$.

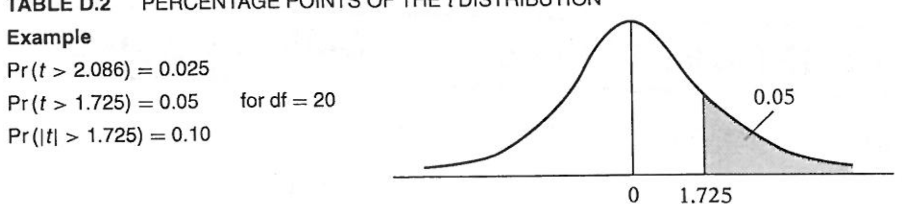
TABLE D.2 PERCENTAGE POINTS OF THE *t* DISTRIBUTION

Example

Pr (t > 2.086) = 0.025

Pr (t > 1.725) = 0.05 for df = 20

Pr (|t| > 1.725) = 0.10



df \ Pr	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Note: The smaller probability shown at the head of each column is the area in one tail; the larger probability is the area in both tails.

Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians,* vol. 1, 3d ed., table 12, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika.*

TABLE D.3 UPPER PERCENTAGE POINTS OF THE *F* DISTRIBUTION

Example

$\Pr(F > 1.59) = 0.25$

$\Pr(F > 2.42) = 0.10$ for df $N_1 = 10$

$\Pr(F > 3.14) = 0.05$ and $N_2 = 9$

$\Pr(F > 5.26) = 0.01$



df for denominator N_2	df for numerator N_1												
	Pr	1	2	3	4	5	6	7	8	9	10	11	12
1	.25	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.36	9.41
	.10	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.5	60.7
	.05	161	200	216	225	230	234	237	239	241	242	243	244
2	.25	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.39
	.10	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.40	9.41
	.05	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4
	.01	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4
3	.25	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.45
	.10	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.22
	.05	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
	.01	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.1
4	.25	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08
	.10	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.91	3.90
	.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
	.01	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.4
5	.25	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89
	.10	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.28	3.27
	.05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.71	4.68
	.01	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.96	9.89
6	.25	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.77
	.10	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.92	2.90
	.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	.01	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
7	.25	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.69	1.69	1.69	1.68
	.10	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.68	2.67
	.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
	.01	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47
8	.25	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.63	1.62
	.10	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.52	2.50
	.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
	.01	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67
9	.25	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
	.10	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.40	2.38
	.05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
	.01	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11

Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians*, vol. 1, 3d ed., table 18, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

df for numerator N_1													df for denominator N_2
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	
9.49	9.58	9.63	9.67	9.71	9.74	9.76	9.78	9.80	9.82	9.84	9.85	.25	1
61.2	61.7	62.0	62.3	62.5	62.7	62.8	63.0	63.1	63.2	63.3	63.3	.10	
246	248	249	250	251	252	252	253	253	254	254	254	.05	
3.41	3.43	3.43	3.44	3.45	3.45	3.46	3.47	3.47	3.48	3.48	3.48	.25	2
9.42	9.44	9.45	9.46	9.47	9.47	9.47	9.48	9.48	9.49	9.49	9.49	.10	
19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	.05	
99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	.01	
2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	.25	3
5.20	5.18	5.18	5.17	5.16	5.15	5.15	5.14	5.14	5.14	5.14	5.13	.10	
8.70	8.66	8.64	8.62	8.59	8.58	8.57	8.55	8.55	8.54	8.53	8.53	.05	
26.9	26.7	26.6	26.5	26.4	26.4	26.3	26.2	26.2	26.2	26.1	26.1	.01	
2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	.25	4
3.87	3.84	3.83	3.82	3.80	3.80	3.79	3.78	3.78	3.77	3.76	3.76	.10	
5.86	5.80	5.77	5.75	5.72	5.70	5.69	5.66	5.66	5.65	5.64	5.63	.05	
14.2	14.0	13.9	13.8	13.7	13.7	13.7	13.6	13.6	13.5	13.5	13.5	.01	
1.89	1.88	1.88	1.88	1.88	1.88	1.87	1.87	1.87	1.87	1.87	1.87	.25	5
3.24	3.21	3.19	3.17	3.16	3.15	3.14	3.13	3.12	3.12	3.11	3.10	.10	
4.62	4.56	4.53	4.50	4.46	4.44	4.43	4.41	4.40	4.39	4.37	4.36	.05	
9.72	9.55	9.47	9.38	9.29	9.24	9.20	9.13	9.11	9.08	9.04	9.02	.01	
1.76	1.76	1.75	1.75	1.75	1.75	1.74	1.74	1.74	1.74	1.74	1.74	.25	6
2.87	2.84	2.82	2.80	2.78	2.77	2.76	2.75	2.74	2.73	2.73	2.72	.10	
3.94	3.87	3.84	3.81	3.77	3.75	3.74	3.71	3.70	3.69	3.68	3.67	.05	
7.56	7.40	7.31	7.23	7.14	7.09	7.06	6.99	6.97	6.93	6.90	6.88	.01	
1.68	1.67	1.67	1.66	1.66	1.66	1.65	1.65	1.65	1.65	1.65	1.65	.25	7
2.63	2.59	2.58	2.56	2.54	2.52	2.51	2.50	2.49	2.48	2.48	2.47	.10	
3.51	3.44	3.41	3.38	3.34	3.32	3.30	3.27	3.27	3.25	3.24	3.23	.05	
6.31	6.16	6.07	5.99	5.91	5.86	5.82	5.75	5.74	5.70	5.67	5.65	.01	
1.62	1.61	1.60	1.60	1.59	1.59	1.59	1.58	1.58	1.58	1.58	1.58	.25	8
2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.32	2.31	2.30	2.29	.10	
3.22	3.15	3.12	3.08	3.04	2.02	3.01	2.97	2.97	2.95	2.94	2.93	.05	
5.52	5.36	5.28	5.20	5.12	5.07	5.03	4.96	4.95	4.91	4.88	4.86	.01	
1.57	1.56	1.56	1.55	1.55	1.54	1.54	1.53	1.53	1.53	1.53	1.53	.25	9
2.34	2.30	2.28	2.25	2.23	2.22	2.21	2.19	2.18	2.17	2.17	2.16	.10	
3.01	2.94	2.90	2.86	2.83	2.80	2.79	2.76	2.75	2.73	2.72	2.71	.05	
4.96	4.81	4.73	4.65	4.57	4.52	4.48	4.42	4.40	4.36	4.33	4.31	.01	

(*Continued*)

TABLE D.3 UPPER PERCENTAGE POINTS OF THE F DISTRIBUTION (*Continued*)

df for denominator N_2	df for numerator N_1												
	Pr	1	2	3	4	5	6	7	8	9	10	11	12
10	.25	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.55	1.54
	.10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.30	2.28
	.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
	.01	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71
11	.25	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.52	1.51
	.10	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.23	2.21
	.05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
	.01	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40
12	.25	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.50	1.49
	.10	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.17	2.15
	.05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
	.01	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16
13	.25	1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47
	.10	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.12	2.10
	.05	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
	.01	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96
14	.25	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.46	1.45
	.10	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.08	2.05
	.05	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
	.01	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
15	.25	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.44
	.10	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.04	2.02
	.05	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
	.01	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
16	.25	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.44	1.43
	.10	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	2.01	1.99
	.05	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42
	.01	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55
17	.25	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.42	1.41
	.10	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.98	1.96
	.05	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38
	.01	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.46
18	.25	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
	.10	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.96	1.93
	.05	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
	.01	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.43	3.37
19	.25	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.40
	.10	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.94	1.91
	.05	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31
	.01	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30
20	.25	1.40	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.39
	.10	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.92	1.89
	.05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28
	.01	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23

df for numerator N_1													df for denominator N_2
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	
1.53	1.52	1.52	1.51	1.51	1.50	1.50	1.49	1.49	1.49	1.48	1.48	.25	10
2.24	2.20	2.18	2.16	2.13	2.12	2.11	2.09	2.08	2.07	2.06	2.06	.10	
2.85	2.77	2.74	2.70	2.66	2.64	2.62	2.59	2.58	2.56	2.55	2.54	.05	
4.56	4.41	4.33	4.25	4.17	4.12	4.08	4.01	4.00	3.96	3.93	3.91	.01	
1.50	1.49	1.49	1.48	1.47	1.47	1.47	1.46	1.46	1.46	1.45	1.45	.25	11
2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	2.00	1.99	1.98	1.97	.10	
2.72	2.65	2.61	2.57	2.53	2.51	2.49	2.46	2.45	2.43	2.42	2.40	.05	
4.25	4.10	4.02	3.94	3.86	3.81	3.78	3.71	3.69	3.66	3.62	3.60	.01	
1.48	1.47	1.46	1.45	1.45	1.44	1.44	1.43	1.43	1.43	1.42	1.42	.25	12
2.10	2.06	2.04	2.01	1.99	1.97	1.96	1.94	1.93	1.92	1.91	1.90	.10	
2.62	2.54	2.51	2.47	2.43	2.40	2.38	2.35	2.34	2.32	2.31	2.30	.05	
4.01	3.86	3.78	3.70	3.62	3.57	3.54	3.47	3.45	3.41	3.38	3.36	.01	
1.46	1.45	1.44	1.43	1.42	1.42	1.42	1.41	1.41	1.40	1.40	1.40	.25	13
2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.88	1.86	1.85	1.85	.10	
2.53	2.46	2.42	2.38	2.34	2.31	2.30	2.26	2.25	2.23	2.22	2.21	.05	
3.82	3.66	3.59	3.51	3.43	3.38	3.34	3.27	3.25	3.22	3.19	3.17	.01	
1.44	1.43	1.42	1.41	1.41	1.40	1.40	1.39	1.39	1.39	1.38	1.38	.25	14
2.01	1.96	1.94	1.91	1.89	1.87	1.86	1.83	1.83	1.82	1.80	1.80	.10	
2.46	2.39	2.35	2.31	2.27	2.24	2.22	2.19	2.18	2.16	2.14	2.13	.05	
3.66	3.51	3.43	3.35	3.27	3.22	3.18	3.11	3.09	3.06	3.03	3.00	.01	
1.43	1.41	1.41	1.40	1.39	1.39	1.38	1.38	1.37	1.37	1.36	1.36	.25	15
1.97	1.92	1.90	1.87	1.85	1.83	1.82	1.79	1.79	1.77	1.76	1.76	.10	
2.40	2.33	2.29	2.25	2.20	2.18	2.16	2.12	2.11	2.10	2.08	2.07	.05	
3.52	3.37	3.29	3.21	3.13	3.08	3.05	2.98	2.96	2.92	2.89	2.87	.01	
1.41	1.40	1.39	1.38	1.37	1.37	1.36	1.36	1.35	1.35	1.34	1.34	.25	16
1.94	1.89	1.87	1.84	1.81	1.79	1.78	1.76	1.75	1.74	1.73	1.72	.10	
2.35	2.28	2.24	2.19	2.15	2.12	2.11	2.07	2.06	2.04	2.02	2.01	.05	
3.41	3.26	3.18	3.10	3.02	2.97	2.93	2.86	2.84	2.81	2.78	2.75	.01	
1.40	1.39	1.38	1.37	1.36	1.35	1.35	1.34	1.34	1.34	1.33	1.33	.25	17
1.91	1.86	1.84	1.81	1.78	1.76	1.75	1.73	1.72	1.71	1.69	1.69	.10	
2.31	2.23	2.19	2.15	2.10	2.08	2.06	2.02	2.01	1.99	1.97	1.96	.05	
3.31	3.16	3.08	3.00	2.92	2.87	2.83	2.76	2.75	2.71	2.68	2.65	.01	
1.39	1.38	1.37	1.36	1.35	1.34	1.34	1.33	1.33	1.32	1.32	1.32	.25	18
1.89	1.84	1.81	1.78	1.75	1.74	1.72	1.70	1.69	1.68	1.67	1.66	.10	
2.27	2.19	2.15	2.11	2.06	2.04	2.02	1.98	1.97	1.95	1.93	1.92	.05	
3.23	3.08	3.00	2.92	2.84	2.78	2.75	2.68	2.66	2.62	2.59	2.57	.01	
1.38	1.37	1.36	1.35	1.34	1.33	1.33	1.32	1.32	1.31	1.31	1.30	.25	19
1.86	1.81	1.79	1.76	1.73	1.71	1.70	1.67	1.67	1.65	1.64	1.63	.10	
2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.94	1.93	1.91	1.89	1.88	.05	
3.15	3.00	2.92	2.84	2.76	2.71	2.67	2.60	2.58	2.55	2.51	2.49	.01	
1.37	1.36	1.35	1.34	1.33	1.33	1.32	1.31	1.31	1.30	1.30	1.29	.25	20
1.84	1.79	1.77	1.74	1.71	1.69	1.68	1.65	1.64	1.63	1.62	1.61	.10	
2.20	2.12	2.08	2.04	1.99	1.97	1.95	1.91	1.90	1.88	1.86	1.84	.05	
3.09	2.94	2.86	2.78	2.69	2.64	2.61	2.54	2.52	2.48	2.44	2.42	.01	

(*Continued*)

TABLE D.3 UPPER PERCENTAGE POINTS OF THE *F* DISTRIBUTION (*Continued*)

df for denominator N_2	df for numerator N_1												
	Pr	1	2	3	4	5	6	7	8	9	10	11	12
22	.25	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.38	1.37
	.10	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86
	.05	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23
	.01	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
24	.25	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.37	1.36
	.10	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83
	.05	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18
	.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
26	.25	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.36	1.35
	.10	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.84	1.81
	.05	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	.01	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
28	.25	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
	.10	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.81	1.79
	.05	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
	.01	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
30	.25	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.35	1.34
	.10	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77
	.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
	.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
40	.25	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31
	.10	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.73	1.71
	.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
	.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
60	.25	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.29
	.10	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66
	.05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
120	.25	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.27	1.26
	.10	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.62	1.60
	.05	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83
	.01	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34
200	.25	1.33	1.39	1.38	1.36	1.34	1.32	1.31	1.29	1.28	1.27	1.26	1.25
	.10	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.57
	.05	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
	.01	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
∞	.25	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.24
	.10	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55
	.05	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75
	.01	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18

df for numerator N_1													df for denominator N_2
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	
1.36	1.34	1.33	1.32	1.31	1.31	1.30	1.30	1.30	1.29	1.29	1.28	.25	22
1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.61	1.60	1.59	1.58	1.57	.10	
2.15	2.07	2.03	1.98	1.94	1.91	1.89	1.85	1.84	1.82	1.80	1.78	.05	
2.98	2.83	2.75	2.67	2.58	2.53	2.50	2.42	2.40	2.36	2.33	2.31	.01	
1.35	1.33	1.32	1.31	1.30	1.29	1.29	1.28	1.28	1.27	1.27	1.26	.25	24
1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.58	1.57	1.56	1.54	1.53	.10	
2.11	2.03	1.98	1.94	1.89	1.86	1.84	1.80	1.79	1.77	1.75	1.73	.05	
2.89	2.74	2.66	2.58	2.49	2.44	2.40	2.33	2.31	2.27	2.24	2.21	.01	
1.34	1.32	1.31	1.30	1.29	1.28	1.28	1.26	1.26	1.26	1.25	1.25	.25	26
1.76	1.71	1.68	1.65	1.61	1.59	1.58	1.55	1.54	1.53	1.51	1.50	.10	
2.07	1.99	1.95	1.90	1.85	1.82	1.80	1.76	1.75	1.73	1.71	1.69	.05	
2.81	2.66	2.58	2.50	2.42	2.36	2.33	2.25	2.23	2.19	2.16	2.13	.01	
1.33	1.31	1.30	1.29	1.28	1.27	1.27	1.26	1.25	1.25	1.24	1.24	.25	28
1.74	1.69	1.66	1.63	1.59	1.57	1.56	1.53	1.52	1.50	1.49	1.48	.10	
2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.73	1.71	1.69	1.67	1.65	.05	
2.75	2.60	2.52	2.44	2.35	2.30	2.26	2.19	2.17	2.13	2.09	2.06	.01	
1.32	1.30	1.29	1.28	1.27	1.26	1.26	1.25	1.24	1.24	1.23	1.23	.25	30
1.72	1.67	1.64	1.61	1.57	1.55	1.54	1.51	1.50	1.48	1.47	1.46	.10	
2.01	1.93	1.89	1.84	1.79	1.76	1.74	1.70	1.68	1.66	1.64	1.62	.05	
2.70	2.55	2.47	2.39	2.30	2.25	2.21	2.13	2.11	2.07	2.03	2.01	.01	
1.30	1.28	1.26	1.25	1.24	1.23	1.22	1.21	1.21	1.20	1.19	1.19	.25	40
1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.43	1.42	1.41	1.39	1.38	.10	
1.92	1.84	1.79	1.74	1.69	1.66	1.64	1.59	1.58	1.55	1.53	1.51	.05	
2.52	2.37	2.29	2.20	2.11	2.06	2.02	1.94	1.92	1.87	1.83	1.80	.01	
1.27	1.25	1.24	1.22	1.21	1.20	1.19	1.17	1.17	1.16	1.15	1.15	.25	60
1.60	1.54	1.51	1.48	1.44	1.41	1.40	1.36	1.35	1.33	1.31	1.29	.10	
1.84	1.75	1.70	1.65	1.59	1.56	1.53	1.48	1.47	1.44	1.41	1.39	.05	
2.35	2.20	2.12	2.03	1.94	1.88	1.84	1.75	1.73	1.68	1.63	1.60	.01	
1.24	1.22	1.21	1.19	1.18	1.17	1.16	1.14	1.13	1.12	1.11	1.10	.25	120
1.55	1.48	1.45	1.41	1.37	1.34	1.32	1.27	1.26	1.24	1.21	1.19	.10	
1.75	1.66	1.61	1.55	1.50	1.46	1.43	1.37	1.35	1.32	1.28	1.25	.05	
2.19	2.03	1.95	1.86	1.76	1.70	1.66	1.56	1.53	1.48	1.42	1.38	.01	
1.23	1.21	1.20	1.18	1.16	1.14	1.12	1.11	1.10	1.09	1.08	1.06	.25	200
1.52	1.46	1.42	1.38	1.34	1.31	1.28	1.24	1.22	1.20	1.17	1.14	.10	
1.72	1.62	1.57	1.52	1.46	1.41	1.39	1.32	1.29	1.26	1.22	1.19	.05	
2.13	1.97	1.89	1.79	1.69	1.63	1.58	1.48	1.44	1.39	1.33	1.28	.01	
1.22	1.19	1.18	1.16	1.14	1.13	1.12	1.09	1.08	1.07	1.04	1.00	.25	∞
1.49	1.42	1.38	1.34	1.30	1.26	1.24	1.18	1.17	1.13	1.08	1.00	.10	
1.67	1.57	1.52	1.46	1.39	1.35	1.32	1.24	1.22	1.17	1.11	1.00	.05	
2.04	1.88	1.79	1.70	1.59	1.52	1.47	1.36	1.32	1.25	1.15	1.00	.01	

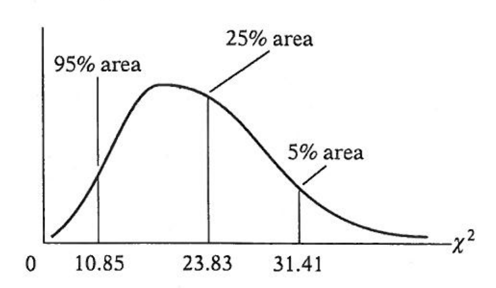
TABLE D.4 UPPER PERCENTAGE POINTS OF THE χ^2 DISTRIBUTION

Example

$\Pr(\chi^2 > 10.85) = 0.95$

$\Pr(\chi^2 > 23.83) = 0.25$ for df = 20

$\Pr(\chi^2 > 31.41) = 0.05$



Degrees of freedom \ Pr	.995	.990	.975	.950	.900
1	392704×10^{-10}	157088×10^{-9}	982069×10^{-9}	393214×10^{-8}	.0157908
2	.0100251	.0201007	.0506356	.102587	.210720
3	.0717212	.114832	.215795	.351846	.584375
4	.206990	.297110	.484419	.710721	1.063623
5	.411740	.554300	.831211	1.145476	1.61031
6	.675727	.872085	1.237347	1.63539	2.20413
7	.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100*	67.3276	70.0648	74.2219	77.9295	82.3581

*For df greater than 100 the expression $\sqrt{2\chi^2} - \sqrt{(2k-1)} = Z$ follows the standardized normal distribution, where k represents the degrees of freedom.

.750	.500	.250	.100	.050	.025	.010	.005
.1015308	.454937	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944
.575364	1.38629	2.77259	4.60517	5.99147	7.37776	9.21034	10.5966
1.212534	2.36597	4.10835	6.25139	7.81473	9.34840	11.3449	12.8381
1.92255	3.35670	5.38527	7.77944	9.48773	11.1433	13.2767	14.8602
2.67460	4.35146	6.62568	9.23635	11.0705	12.8325	15.0863	16.7496
3.45460	5.34812	7.84080	10.6446	12.5916	14.4494	16.8119	18.5476
4.25485	6.34581	9.03715	12.0170	14.0671	16.0128	18.4753	20.2777
5.07064	7.34412	10.2188	13.3616	15.5073	17.5346	20.0902	21.9550
5.89883	8.34283	11.3887	14.6837	16.9190	19.0228	21.6660	23.5893
6.73720	9.34182	12.5489	15.9871	18.3070	20.4831	23.2093	25.1882
7.58412	10.3410	13.7007	17.2750	19.6751	21.9200	24.7250	26.7569
8.43842	11.3403	14.8454	18.5494	21.0261	23.3367	26.2170	28.2995
9.29906	12.3398	15.9839	19.8119	22.3621	24.7356	27.6883	29.8194
10.1653	13.3393	17.1170	21.0642	23.6848	26.1190	29.1413	31.3193
11.0365	14.3389	18.2451	22.3072	24.9958	27.4884	30.5779	32.8013
11.9122	15.3385	19.3688	23.5418	26.2962	28.8454	31.9999	34.2672
12.7919	16.3381	20.4887	24.7690	27.5871	30.1910	33.4087	35.7185
13.6753	17.3379	21.6049	25.9894	28.8693	31.5264	34.8053	37.1564
14.5620	18.3376	22.7178	27.2036	30.1435	32.8523	36.1908	38.5822
15.4518	19.3374	23.8277	28.4120	31.4104	34.1696	37.5662	39.9968
16.3444	20.3372	24.9348	29.6151	32.6705	35.4789	38.9321	41.4010
17.2396	21.3370	26.0393	30.8133	33.9244	36.7807	40.2894	42.7956
18.1373	22.3369	27.1413	32.0069	35.1725	38.0757	41.6384	44.1813
19.0372	23.3367	28.2412	33.1963	36.4151	39.3641	42.9798	45.5585
19.9393	24.3366	29.3389	34.3816	37.6525	40.6465	44.3141	46.9278
20.8434	25.3364	30.4345	35.5631	38.8852	41.9232	45.6417	48.2899
21.7494	26.3363	31.5284	36.7412	40.1133	43.1944	46.9630	49.6449
22.6572	27.3363	32.6205	37.9159	41.3372	44.4607	48.2782	50.9933
23.5666	28.3362	33.7109	39.0875	42.5569	45.7222	49.5879	52.3356
24.4776	29.3360	34.7998	40.2560	43.7729	46.9792	50.8922	53.6720
33.6603	39.3354	45.6160	51.8050	55.7585	59.3417	63.6907	66.7659
42.9421	49.3349	56.3336	63.1671	67.5048	71.4202	76.1539	79.4900
52.2938	59.3347	66.9814	74.3970	79.0819	83.2976	88.3794	91.9517
61.6983	69.3344	77.5766	85.5271	90.5312	95.0231	100.425	104.215
71.1445	79.3343	88.1303	96.5782	101.879	106.629	112.329	116.321
80.6247	89.3342	98.6499	107.565	113.145	118.136	124.116	128.299
90.1332	99.3341	109.141	118.498	124.342	129.561	135.807	140.169

Source: Abridged from E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians*, vol. 1, 3d ed., table 8, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

TABLE D.5A

DURBIN–WATSON *d* STATISTIC: SIGNIFICANCE POINTS OF d_L AND d_U AT 0.05 LEVEL OF SIGNIFICANCE

	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$		$k'=8$		$k'=9$		$k'=10$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—	—	—	—	—	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—	—	—	—	—	—	—
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005	—	—	—	—	—	—	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149	—	—	—	—	—	—
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266	—	—	—	—
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	—	—
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	2.071	0.798	2.188	0.723	2.309	0.650	2.431
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.682	2.396
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.795	2.281
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	1.979	0.950	2.069	0.885	2.162	0.821	2.257
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.198
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.149
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874

	$k'=11$		$k'=12$		$k'=13$		$k'=14$		$k'=15$		$k'=16$		$k'=17$		$k'=18$		$k'=19$		$k'=20$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0.098	3.503	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.138	3.378	0.087	3.557	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.177	3.265	0.123	3.441	0.078	3.603	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.220	3.159	0.160	3.335	0.111	3.496	0.070	3.642	—	—	—	—	—	—	—	—	—	—	—	—
20	0.263	3.063	0.200	3.234	0.145	3.395	0.100	3.542	0.063	3.676	—	—	—	—	—	—	—	—	—	—
21	0.307	2.976	0.240	3.141	0.182	3.300	0.132	3.448	0.091	3.583	0.058	3.705	—	—	—	—	—	—	—	—
22	0.349	2.897	0.281	3.057	0.220	3.211	0.166	3.358	0.120	3.495	0.083	3.619	0.052	3.731	—	—	—	—	—	—
23	0.391	2.826	0.322	2.979	0.259	3.128	0.202	3.272	0.153	3.409	0.110	3.535	0.076	3.650	0.048	3.753	—	—	—	—
24	0.431	2.761	0.362	2.908	0.297	3.053	0.239	3.193	0.186	3.327	0.141	3.454	0.101	3.572	0.070	3.678	0.044	3.773	—	—
25	0.470	2.702	0.400	2.844	0.335	2.983	0.275	3.119	0.221	3.251	0.172	3.376	0.130	3.494	0.094	3.604	0.065	3.702	0.041	3.790
26	0.508	2.649	0.438	2.784	0.373	2.919	0.312	3.051	0.256	3.179	0.205	3.303	0.160	3.420	0.120	3.531	0.087	3.632	0.060	3.724
27	0.544	2.600	0.475	2.730	0.409	2.859	0.348	2.987	0.291	3.112	0.238	3.233	0.191	3.349	0.149	3.460	0.112	3.563	0.081	3.658
28	0.578	2.555	0.510	2.680	0.445	2.805	0.383	2.928	0.325	3.050	0.271	3.168	0.222	3.283	0.178	3.392	0.138	3.495	0.104	3.592
29	0.612	2.515	0.544	2.634	0.479	2.755	0.418	2.874	0.359	2.992	0.305	3.107	0.254	3.219	0.208	3.327	0.166	3.431	0.129	3.528
30	0.643	2.477	0.577	2.592	0.512	2.708	0.451	2.823	0.392	2.937	0.337	3.050	0.286	3.160	0.238	3.266	0.195	3.368	0.156	3.465
31	0.674	2.443	0.608	2.553	0.545	2.665	0.484	2.776	0.425	2.887	0.370	2.996	0.317	3.103	0.269	3.208	0.224	3.309	0.183	3.406
32	0.703	2.411	0.638	2.517	0.576	2.625	0.515	2.733	0.457	2.840	0.401	2.946	0.349	3.050	0.299	3.153	0.253	3.252	0.211	3.348
33	0.731	2.382	0.668	2.484	0.606	2.588	0.546	2.692	0.488	2.796	0.432	2.899	0.379	3.000	0.329	3.100	0.283	3.198	0.239	3.293
34	0.758	2.355	0.695	2.454	0.634	2.554	0.575	2.654	0.518	2.754	0.462	2.854	0.409	2.954	0.359	3.051	0.312	3.147	0.267	3.240
35	0.783	2.330	0.722	2.425	0.662	2.521	0.604	2.619	0.547	2.716	0.492	2.813	0.439	2.910	0.388	3.005	0.340	3.099	0.295	3.190
36	0.808	2.306	0.748	2.398	0.689	2.492	0.631	2.586	0.575	2.680	0.520	2.774	0.467	2.868	0.417	2.961	0.369	3.053	0.323	3.142
37	0.831	2.285	0.772	2.374	0.714	2.464	0.657	2.555	0.602	2.646	0.548	2.738	0.495	2.829	0.445	2.920	0.397	3.009	0.351	3.097
38	0.854	2.265	0.796	2.351	0.739	2.438	0.683	2.526	0.628	2.614	0.575	2.703	0.522	2.792	0.472	2.880	0.424	2.968	0.378	3.054
39	0.875	2.246	0.819	2.329	0.763	2.413	0.707	2.499	0.653	2.585	0.600	2.671	0.549	2.757	0.499	2.843	0.451	2.929	0.404	3.013
40	0.896	2.228	0.840	2.309	0.785	2.391	0.731	2.473	0.678	2.557	0.626	2.641	0.575	2.724	0.525	2.808	0.477	2.892	0.430	2.974
45	0.988	2.156	0.938	2.225	0.887	2.296	0.838	2.367	0.788	2.439	0.740	2.512	0.692	2.586	0.644	2.659	0.598	2.733	0.553	2.807
50	1.064	2.103	1.019	2.163	0.973	2.225	0.927	2.287	0.882	2.350	0.836	2.414	0.792	2.479	0.747	2.544	0.703	2.610	0.660	2.675
55	1.129	2.062	1.087	2.116	1.045	2.170	1.003	2.225	0.961	2.281	0.919	2.338	0.877	2.396	0.836	2.454	0.795	2.512	0.754	2.571
60	1.184	2.031	1.145	2.079	1.106	2.127	1.068	2.177	1.029	2.227	0.990	2.278	0.951	2.330	0.913	2.382	0.874	2.434	0.836	2.487
65	1.231	2.006	1.195	2.049	1.160	2.093	1.124	2.138	1.088	2.183	1.052	2.229	1.016	2.276	0.980	2.323	0.944	2.371	0.908	2.419
70	1.272	1.986	1.239	2.026	1.206	2.066	1.172	2.106	1.139	2.148	1.105	2.189	1.072	2.232	1.038	2.275	1.005	2.318	0.971	2.362
75	1.308	1.970	1.277	2.006	1.247	2.043	1.215	2.080	1.184	2.118	1.153	2.156	1.121	2.195	1.090	2.235	1.058	2.275	1.027	2.315
80	1.340	1.957	1.311	1.991	1.283	2.024	1.253	2.059	1.224	2.093	1.195	2.129	1.165	2.165	1.136	2.201	1.106	2.238	1.076	2.275
85	1.369	1.946	1.342	1.977	1.315	2.009	1.287	2.040	1.260	2.073	1.232	2.105	1.205	2.139	1.177	2.172	1.149	2.206	1.121	2.241
90	1.395	1.937	1.369	1.966	1.344	1.995	1.318	2.025	1.292	2.055	1.266	2.085	1.240	2.116	1.213	2.148	1.187	2.179	1.160	2.211
95	1.418	1.929	1.394	1.956	1.370	1.984	1.345	2.012	1.321	2.040	1.296	2.068	1.271	2.097	1.247	2.126	1.222	2.156	1.197	2.186
100	1.439	1.923	1.416	1.948	1.393	1.974	1.371	2.000	1.347	2.026	1.324	2.053	1.301	2.080	1.277	2.108	1.253	2.135	1.229	2.164
150	1.579	1.892	1.564	1.908	1.550	1.924	1.535	1.940	1.519	1.956	1.504	1.972	1.489	1.989	1.474	2.006	1.458	2.023	1.443	2.040
200	1.654	1.885	1.643	1.896	1.632	1.908	1.621	1.919	1.610	1.931	1.599	1.943	1.588	1.955	1.576	1.967	1.565	1.979	1.554	1.991

Source: This table is an extension of the original Durbin–Watson table and is reproduced from N. E. Savin and K. J. White, "The Durbin-Watson Test for Serial Correlation with Extreme Small Samples or Many Regressors," *Econometrica*, vol. 45, November 1977, pp. 1989–96 and as corrected by R. W. Farebrother, *Econometrica*, vol. 48, September 1980, p. 1554. Reprinted by permission of the Econometric Society.

Note: n = number of observations, k' = number of explanatory variables excluding the constant term.

EXAMPLE

If $n = 40$ and $k' = 4$, $d_L = 1.285$ and $d_U = 1.721$. If a computed d value is less than 1.285, there is evidence of positive first-order serial correlation; if it is greater than 1.721, there is no evidence of positive first-order serial correlation; but if d lies between the lower and the upper limit, there is inconclusive evidence regarding the presence or absence of positive first-order serial correlation.

TABLE D.5B

DURBIN–WATSON *d* STATISTIC: SIGNIFICANCE POINTS OF d_L AND d_U AT 0.01 LEVEL OF SIGNIFICANCE

	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$		$k'=8$		$k'=9$		$k'=10$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.390	1.142	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.435	1.036	0.294	1.676	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.497	1.003	0.345	1.489	0.229	2.102	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.554	0.998	0.408	1.389	0.279	1.875	0.183	2.433	—	—	—	—	—	—	—	—	—	—	—	—
10	0.604	1.001	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690	—	—	—	—	—	—	—	—	—	—
11	0.653	1.010	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892	—	—	—	—	—	—	—	—
12	0.697	1.023	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053	—	—	—	—	—	—
13	0.738	1.038	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182	—	—	—	—
14	0.776	1.054	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2.354	0.183	2.667	0.122	2.981	0.078	3.287	—	—
15	0.811	1.070	0.700	1.252	0.591	1.464	0.488	1.704	0.391	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374
16	0.844	1.086	0.737	1.252	0.633	1.446	0.532	1.663	0.437	1.900	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944	0.094	3.201
17	0.874	1.102	0.772	1.255	0.672	1.432	0.574	1.630	0.480	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053
18	0.902	1.118	0.805	1.259	0.708	1.422	0.613	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160	2.925
19	0.928	1.132	0.835	1.265	0.742	1.415	0.650	1.584	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196	2.813
20	0.952	1.147	0.863	1.271	0.773	1.411	0.685	1.567	0.598	1.737	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232	2.714
21	0.975	1.161	0.890	1.277	0.803	1.408	0.718	1.554	0.633	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268	2.625
22	0.997	1.174	0.914	1.284	0.831	1.407	0.748	1.543	0.667	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304	2.548
23	1.018	1.187	0.938	1.291	0.858	1.407	0.777	1.534	0.698	1.673	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.340	2.479
24	1.037	1.199	0.960	1.298	0.882	1.407	0.805	1.528	0.728	1.658	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375	2.417
25	1.055	1.211	0.981	1.305	0.906	1.409	0.831	1.523	0.756	1.645	0.682	1.776	0.610	1.915	0.540	2.059	0.473	2.209	0.409	2.362
26	1.072	1.222	1.001	1.312	0.928	1.411	0.855	1.518	0.783	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441	2.313
27	1.089	1.233	1.019	1.319	0.949	1.413	0.878	1.515	0.808	1.626	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473	2.269
28	1.104	1.244	1.037	1.325	0.969	1.415	0.900	1.513	0.832	1.618	0.764	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504	2.229
29	1.119	1.254	1.054	1.332	0.988	1.418	0.921	1.512	0.855	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533	2.193
30	1.133	1.263	1.070	1.339	1.006	1.421	0.941	1.511	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160
31	1.147	1.273	1.085	1.345	1.023	1.425	0.960	1.510	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589	2.131
32	1.160	1.282	1.100	1.352	1.040	1.428	0.979	1.510	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615	2.104
33	1.172	1.291	1.114	1.358	1.055	1.432	0.996	1.510	0.936	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641	2.080
34	1.184	1.299	1.128	1.364	1.070	1.435	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665	2.057
35	1.195	1.307	1.140	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689	2.037
36	1.206	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.988	1.588	0.932	1.666	0.877	1.749	0.821	1.836	0.766	1.925	0.711	2.018
37	1.217	1.323	1.165	1.382	1.112	1.446	1.058	1.514	1.004	1.586	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733	2.001
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.585	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754	1.985
39	1.237	1.337	1.187	1.393	1.137	1.453	1.085	1.517	1.034	1.584	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774	1.970
40	1.246	1.344	1.198	1.398	1.148	1.457	1.098	1.518	1.048	1.584	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.749	1.956
45	1.288	1.376	1.245	1.423	1.201	1.474	1.156	1.528	1.111	1.584	1.065	1.643	1.019	1.704	0.974	1.768	0.927	1.834	0.881	1.902
50	1.324	1.403	1.285	1.446	1.245	1.491	1.205	1.538	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955	1.864
55	1.356	1.427	1.320	1.466	1.284	1.506	1.247	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.734	1.057	1.785	1.018	1.837
60	1.383	1.449	1.350	1.484	1.317	1.520	1.283	1.558	1.249	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072	1.817
65	1.407	1.468	1.377	1.500	1.346	1.534	1.315	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120	1.802
70	1.429	1.485	1.400	1.515	1.372	1.546	1.343	1.578	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162	1.792
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.587	1.340	1.617	1.313	1.649	1.284	1.682	1.256	1.714	1.227	1.748	1.199	1.783
80	1.466	1.515	1.441	1.541	1.416	1.568	1.390	1.595	1.364	1.624	1.338	1.653	1.312	1.683	1.285	1.714	1.259	1.745	1.232	1.777
85	1.482	1.528	1.458	1.553	1.435	1.578	1.411	1.603	1.386	1.630	1.362	1.657	1.337	1.685	1.312	1.714	1.287	1.743	1.262	1.773
90	1.496	1.540	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.360	1.687	1.336	1.714	1.312	1.741	1.288	1.769
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.642	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313	1.767
100	1.522	1.562	1.503	1.583	1.482	1.604	1.462	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335	1.765
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486	1.767
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.757	1.582	1.768	1.571	1.779

	k′ = 11		k′ = 12		k′ = 13		k′ = 14		k′ = 15		k′ = 16		k′ = 17		k′ = 18		k′ = 19		k′ = 20	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0.060	3.446	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.084	3.286	0.053	3.506	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.113	3.146	0.075	3.358	0.047	3.357	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.145	3.023	0.102	3.227	0.067	3.420	0.043	3.601	—	—	—	—	—	—	—	—	—	—	—	—
20	0.178	2.914	0.131	3.109	0.092	3.297	0.061	3.474	0.038	3.639	—	—	—	—	—	—	—	—	—	—
21	0.212	2.817	0.162	3.004	0.119	3.185	0.084	3.358	0.055	3.521	0.035	3.671	—	—	—	—	—	—	—	—
22	0.246	2.729	0.194	2.909	0.148	3.084	0.109	3.252	0.077	3.412	0.050	3.562	0.032	3.700	—	—	—	—	—	—
23	0.281	2.651	0.227	2.822	0.178	2.991	0.136	3.155	0.100	3.311	0.070	3.459	0.046	3.597	0.029	3.725	—	—	—	—
24	0.315	2.580	0.260	2.744	0.209	2.906	0.165	3.065	0.125	3.218	0.092	3.363	0.065	3.501	0.043	3.629	0.027	3.747	—	—
25	0.348	2.517	0.292	2.674	0.240	2.829	0.194	2.982	0.152	3.131	0.116	3.274	0.085	3.410	0.060	3.538	0.039	3.657	0.025	3.766
26	0.381	2.460	0.324	2.610	0.272	2.758	0.224	2.906	0.180	3.050	0.141	3.191	0.107	3.325	0.079	3.452	0.055	3.572	0.036	3.682
27	0.413	2.409	0.356	2.552	0.303	2.694	0.253	2.836	0.208	2.976	0.167	3.113	0.131	3.245	0.100	3.371	0.073	3.490	0.051	3.602
28	0.444	2.363	0.387	2.499	0.333	2.635	0.283	2.772	0.237	2.907	0.194	3.040	0.156	3.169	0.122	3.294	0.093	3.412	0.068	3.524
29	0.474	2.321	0.417	2.451	0.363	2.582	0.313	2.713	0.266	2.843	0.222	2.972	0.182	3.098	0.146	3.220	0.114	3.338	0.087	3.450
30	0.503	2.283	0.447	2.407	0.393	2.533	0.342	2.659	0.294	2.785	0.249	2.909	0.208	3.032	0.171	3.152	0.137	3.267	0.107	3.379
31	0.531	2.248	0.475	2.367	0.422	2.487	0.371	2.609	0.322	2.730	0.277	2.851	0.234	2.970	0.196	3.087	0.160	3.201	0.128	3.311
32	0.558	2.216	0.503	2.330	0.450	2.446	0.399	2.563	0.350	2.680	0.304	2.797	0.261	2.912	0.221	3.026	0.184	3.137	0.151	3.246
33	0.585	2.187	0.530	2.296	0.477	2.408	0.426	2.520	0.377	2.633	0.331	2.746	0.287	2.858	0.246	2.969	0.209	3.078	0.174	3.184
34	0.610	2.160	0.556	2.266	0.503	2.373	0.452	2.481	0.404	2.590	0.357	2.699	0.313	2.808	0.272	2.915	0.233	3.022	0.197	3.126
35	0.634	2.136	0.581	2.237	0.529	2.340	0.478	2.444	0.430	2.550	0.383	2.655	0.339	2.761	0.297	2.865	0.257	2.969	0.221	3.071
36	0.658	2.113	0.605	2.210	0.554	2.310	0.504	2.410	0.455	2.512	0.409	2.614	0.364	2.717	0.322	2.818	0.282	2.919	0.244	3.019
37	0.680	2.092	0.628	2.186	0.578	2.282	0.528	2.379	0.480	2.477	0.434	2.576	0.389	2.675	0.347	2.774	0.306	2.872	0.268	2.969
38	0.702	2.073	0.651	2.164	0.601	2.256	0.552	2.350	0.504	2.445	0.458	2.540	0.414	2.637	0.371	2.733	0.330	2.828	0.291	2.923
39	0.723	2.055	0.673	2.143	0.623	2.232	0.575	2.323	0.528	2.414	0.482	2.507	0.438	2.600	0.395	2.694	0.354	2.787	0.315	2.879
40	0.744	2.039	0.694	2.123	0.645	2.210	0.597	2.297	0.551	2.386	0.505	2.476	0.461	2.566	0.418	2.657	0.377	2.748	0.338	2.838
45	0.835	1.972	0.790	2.044	0.744	2.118	0.700	2.193	0.655	2.269	0.612	2.346	0.570	2.424	0.528	2.503	0.488	2.582	0.448	2.661
50	0.913	1.925	0.871	1.987	0.829	2.051	0.787	2.116	0.746	2.182	0.705	2.250	0.665	2.318	0.625	2.387	0.586	2.456	0.548	2.526
55	0.979	1.891	0.940	1.945	0.902	2.002	0.863	2.059	0.825	2.117	0.786	2.176	0.748	2.237	0.711	2.298	0.674	2.359	0.637	2.421
60	1.037	1.865	1.001	1.914	0.965	1.964	0.929	2.015	0.893	2.067	0.857	2.120	0.822	2.173	0.786	2.227	0.751	2.283	0.716	2.338
65	1.087	1.845	1.053	1.889	1.020	1.934	0.986	1.980	0.953	2.027	0.919	2.075	0.886	2.123	0.852	2.172	0.819	2.221	0.786	2.272
70	1.131	1.831	1.099	1.870	1.068	1.911	1.037	1.953	1.005	1.995	0.974	2.038	0.943	2.082	0.911	2.127	0.880	2.172	0.849	2.217
75	1.170	1.819	1.141	1.856	1.111	1.893	1.082	1.931	1.052	1.970	1.023	2.009	0.993	2.049	0.964	2.090	0.934	2.131	0.905	2.172
80	1.205	1.810	1.177	1.844	1.150	1.878	1.122	1.913	1.094	1.949	1.066	1.984	1.039	2.022	1.011	2.059	0.983	2.097	0.955	2.135
85	1.236	1.803	1.210	1.834	1.184	1.866	1.158	1.898	1.132	1.931	1.106	1.965	1.080	1.999	1.053	2.033	1.027	2.068	1.000	2.104
90	1.264	1.798	1.240	1.827	1.215	1.856	1.191	1.886	1.166	1.917	1.141	1.948	1.116	1.979	1.091	2.012	1.066	2.044	1.041	2.077
95	1.290	1.793	1.267	1.821	1.244	1.848	1.221	1.876	1.197	1.905	1.174	1.934	1.150	1.963	1.126	1.993	1.102	2.023	1.079	2.054
100	1.314	1.790	1.292	1.816	1.270	1.841	1.248	1.868	1.225	1.895	1.203	1.922	1.181	1.949	1.158	1.977	1.136	2.006	1.113	2.034
150	1.473	1.783	1.458	1.799	1.444	1.814	1.429	1.830	1.414	1.847	1.400	1.863	1.385	1.880	1.370	1.897	1.355	1.913	1.340	1.931
200	1.561	1.791	1.550	1.801	1.539	1.813	1.528	1.824	1.518	1.836	1.507	1.847	1.495	1.860	1.484	1.871	1.474	1.883	1.462	1.896

Note: n = number of observations

k' = number of explanatory variables excluding the constant term.

Source: Savin and White, op. cit., by permission of the Econometric Society.

TABLE D.6A CRITICAL VALUES OF RUNS IN THE RUNS TEST

	N_2																		
N_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

Note: Tables D.6A and D.6B give the critical values of runs *n* for various values of N_1 (+ symbol) and N_2 (– symbol). For the one-sample runs test, any value of *n* that is equal to or smaller than that shown in Table D.6A or equal to or larger than that shown in Table D.6B is significant at the 0.05 level.

Source: Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, New York, 1956, table F, pp. 252–253. The tables have been adapted by Siegel from the original source: Frieda S. Swed and C. Eisenhart, "Tables for Testing Randomness of Grouping in a Sequence of Alternatives," *Annals of Mathematical Statistics*, vol. 14, 1943. Used by permission of McGraw-Hill Book Company and *Annals of Mathematical Statistics*.

TABLE D.6B CRITICAL VALUES OF RUNS IN THE RUNS TEST

	N_2																		
N_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

EXAMPLE

In a sequence of 30 observations consisting of 20 + signs ($= N_1$) and 10 − signs ($= N_2$), the critical values of runs at the 0.05 level of significance are 9 and 20, as shown by Tables D.6A and D.6B, respectively. Therefore, if in an application it is found that the number of runs is equal to or less than 9 or equal to or greater than 20, one can reject (at the 0.05 level of significance) the hypothesis that the observed sequence is random.

TABLE D.7 1% AND 5% CRITICAL DICKEY–FULLER t ($= \tau$) AND F VALUES FOR UNIT ROOT TESTS

Sample size	t_{nc}*		t_c*		t_{ct}*		F†		F‡	
	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%
25	−2.66	−1.95	−3.75	−3.00	−4.38	−3.60	10.61	7.24	8.21	5.68
50	−2.62	−1.95	−3.58	−2.93	−4.15	−3.50	9.31	6.73	7.02	5.13
100	−2.60	−1.95	−3.51	−2.89	−4.04	−3.45	8.73	6.49	6.50	4.88
250	−2.58	−1.95	−3.46	−2.88	−3.99	−3.43	8.43	6.34	6.22	4.75
500	−2.58	−1.95	−3.44	−2.87	−3.98	−3.42	8.34	6.30	6.15	4.71
∞	−2.58	−1.95	−3.43	−2.86	−3.96	−3.41	8.27	6.25	6.09	4.68

*Subscripts nc, c, and ct denote, respectively, that there is no constant, a constant, and a constant and trend term in the regression (21.9.5).

†The critical F values are for the joint hypothesis that the constant and δ terms in (21.9.5) are simultaneously equal to zero.

‡The critical F values are for the joint hypothesis that the constant, trend, and δ terms in (21.9.5) are simultaneously equal to zero.

Source: Adapted from W. A. Fuller, *Introduction to Statistical Time Series*, John Wiley & Sons, New York, 1976, p. 373 (for the τ test), and D.A. Dickey and W. A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, vol. 49, 1981, p. 1063.